THE UNIVERSITY OF CHICAGO

INTEGRATING MOLLUSCAN DEAD-SHELL ASSEMBLAGES WITH LONG-TERM BIOLOGICAL MONITORING: ARCHIVES OF SPATIO-TEMPORAL GRADIENTS IN PUGET SOUND, SOUTHERN CALIFORNIA, AND JAMAICA

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF THE GEOPHYSICAL SCIENCES

BY
BROC STEPHEN KOKESH

CHICAGO, ILLINOIS
MARCH 2024

For my parents

I have long been a student of sequences,

probably because of my upbringing on a farm...

...you simply don't go out and do a piece of work.

No, the first thing you do is determine the lengthy sequence

of activities necessary even to begin the job.

Then you realize that the sequence of preparatory activities

is so long you will never get to the intended task.

So you go fishing instead.

- *Patrick F. McManus*

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I first wish to thank my primary advisor – Susan Kidwell – for her mentorship and unwavering support these past few years. Sue, words cannot express my gratitude. From brainstorming with me on the lunchroom chalkboard, highlighting my many drafts of the papers we have written, and discussing what it is I want to be when I grow up, you've given me the time and patience I could only hope to give future students of my own. I learn something new and seemingly devise a new project each time we meet. This document could easily have two or three extra chapters with all of our current works in progress, but I'll instead look forward to actualizing those with you in the future.

To the members of my committee – Michael Foote, David Jablonski, and Scott Lidgard – thank you for every patiently scheduled discussion, every thoughtful interjection of new ideas, and every critique of where my science has stumbled. Michael, thank you for pushing me to understand the nature of my data, analyses, and all the wonderful output they generate. I now stop and think before every correlation test about which variable is truly dependent on the other. Dave, your passion for bivalves is infectious and you never cease to help me see the larger implications. Participating in your discussion courses and lab group meetings has forever shaped the way I hope to approach teaching and mentoring others. Scott, I always enjoy your reminders to clearly define and emphasize the motivating questions. You've been an invaluable member of this team and I can't thank you enough for signing on from the very first inquiry.

I thank the current and former members of the Washington State Department of Ecology's Marine Monitoring Unit for their extensive and ongoing collaboration in Puget Sound: Margaret Dutch, Dany Burgess, Sandra Weakland, Valerie Partritch, Angela Eagleston, and Micah Horwith. I look forward to producing the other exciting projects we have in the pipeline together and continuing to archive the "grunge" from your samples.

I thank the scientists and staff at the Los Angeles County Sanitation Districts (LACSD),

especially Shelly Walther and Don Cadien, for their collaboration and support. Thanks also to the dedicated community of the Southern California Association of Marine Invertebrate Taxonomists (SCAMIT) for maintaining and updating their exceptional list of benthic invertebrate taxa referenced herein over the decades. Chapters 2 and 3 are part of Special Project JWSS-19-003 of the LACSD, approved by the Los Angeles Regional Water Quality Control Board.

I thank everyone with the Department of Geography and Geology at the University of the West Indies for making Jamaica a second home. Thank you to Thomas Stemann for serving as my project sponsor. Thank you to Hugh Small and the staff of the Port Royal Marine Laboratory for their assistance in coordination field work and the US Embassy in Kingston for logistical support. Thank you to Dayne Buddo, Mona Webber, Suzanne Palmer, and Michael Burn for insightful conversations about the current and historical state of the Kingston Harbour.

Thank you to my fellow students, colleagues, and friends across the Geophysical Sciences, E&E, OBA, and CEB programs, especially those on the second-floor of Hinds. To Rachel Laker (my academic twin), thank you for sharing this journey from our interview weekend to accepting our postdocs! To Reuben Ng and Melissa Wood, thanks for making many a weekend sociable at a time when the world required smaller circles. To the DOGS softball team and Liquidus regulars, thank you for capping the work week off with fun and camaraderie – Cheers! Thank you to Adam Tomašových for insightful conversations on data and coding. To the undergraduate students who felt the itch to pick shells and passed through our lab – Ben Shafer, Lilja Carden, Thomas Cortelessi, Maskin Sidhu, and Clair Long – thank you for your contributions, and I wish you all the best.

Thank you to my family and loved ones for constantly supporting my journey and anchoring me to reality when science could wait: my mother Kaija; siblings Kassie, Wyatt, and Kate; grandparents; cousins; in-laws; and everyone else I'd fill pages to name. To my father

Brian, not a day goes by where I don't think about you and miss your presence; You truly were the catalyst that sparked my passion for science, and I dedicate the work produced in this document to you. Finally, thank you to my wife, Shannon, for your unwavering love, understanding, and companionship; I can't wait for the next adventure!

Funding for the projects detailed in this document was provided by a New Direction grant from the Petroleum Research Fund of the American Chemical Society (PRF58795-ND8), the University of Southern California Sea Grant Program, the National Oceanic and Atmospheric Administration (NA07OAR4170008), the National Science Foundation (EAR-1124189), and a Student Research Award through the US Fulbright Program.

ABSTRACT

'Live-dead comparison' refers to the measurement of compositional fidelity between living communities and co-occurring accumulations of dead skeletal material (time-averaged 'death assemblages') and has long been used as a means of validating taphonomic and paleoecological interpretations in the fossil record. More recently, live-dead comparisons have been applied as a tool for detecting biological strain resulting from anthropogenic stressors on local to regional scales. However, two important limitations to live-dead comparisons remain underexplored: (1) to what extent are the organisms that readily-preserve in death assemblages (and the fossil record) a sufficient proxy for the broader ecological patterns of living communities that include a plurality of soft-bodied clades, and (2) what temporal context and variation is lost when – as is typically the case – the living community is only sampled once alongside the dead (i.e., a single 'snapshot' of standing diversity)? Examining these questions requires time series of living community compositional data with a broad taxonomic scope; such datasets are the purview of long-term biomonitoring programs, be they government-mandated, contracted, or non-profit. The following chapters represent collaborative efforts between academic paleontologists and agency ecologists to assess the power of bivalve death assemblages as surrogates of the whole-fauna community dynamics using exceptionally-long time series or historical records in three environmentally-distinct locations: (1) the naturally-variable regional benthos of Puget Sound (Chapter 1; 30-year living time series + death assemblage), (2) the historically-polluted Palos Verdes shelf off of Los Angeles County (Chapters 2-3; 50-year living time series + two death assemblages), and (3) the mangrove swamps of Jamaica's Kingston Harbour (Chapter 4; one live-dead sampling + historic surveys of the living).

Surrogates of Compositional Variation. To integrate paleoecological data with the 'whole fauna' data used in biological monitoring, analyses must usually focus on datasets that meet the same standards of numerical and taxonomic detail and scope. Here, the ability

of readily-preserved bivalves to reflect patterns of compositional variation from the entire macroinvertebrate fauna was assessed using data from ten long-established subtidal stations in Puget Sound, Washington State. Similarity in compositional variation was assessed for five taxonomic subsets (the whole fauna, polychaetes, malacostracans, living bivalves, dead bivalves) at four levels of taxonomic resolution (species, genera, families, orders) evaluated under four numerical transformations of the original count data (proportional abundance, square root- and fourth root-transformation, presence-absence). Using the original matrix of species-level proportional abundances of the whole fauna as a benchmark of 'compositional variation,' living and dead bivalves had nearly identical potential to serve as surrogates of the whole fauna; they were further offset than the polychaete subset, but far superior as surrogates than malacostracans. Genus- and family-level data were consistently strong surrogates of species-level data, and the impacts of data transformation were closely tied to taxonomic evenness. The strong congruence of death assemblages with living bivalves is encouraging for using bivalve dead-shell assemblages to complement conventional monitoring data, notwithstanding strong natural environmental gradients with potential to bias shell preservation.

Detectors of Historic Pollution Gradients. Surrogates of macrobenthic assemblages can take many forms, such as using coarser taxonomic levels ('sufficiency') or only a subset of the whole fauna ('subsetting'). Here, the power of both approaches to retain community-level patterns of spatial and temporal variation were evaluated using an exceptionally long (47-year) infaunal dataset generated from monitoring wastewater impacts on an urban shelf in southern California. Four taxonomic sets (whole infauna, polychaetes, bivalves, malacostracans) were evaluated at five resolutions (species, genus, family, order, functional guild) along a pollution gradient subdivided into two spatial bins based on proximity to the wastewater outfall (near-field vs far-field) and three temporal bins based on wastewater treatment phases. All taxonomic sets detected weakening of the spatial gradient with improved wastew-

ater treatment – communities became more similar in richness, evenness, and composition through time – and patterns were robust when coarsened to families or guilds. Polychaetes mirrored ('proxied') whole-fauna patterns most accurately, but bivalves outperformed all other sets in detecting ('indicating') the pollution gradient itself. As taxonomic surrogates, the selection of a subset to represent the whole requires consideration of each clade's strengths to serve different monitoring objectives.

Indicators of Ecological Quality Status. Death assemblages are promising tools to evaluate ecological quality from the past owing to their time-averaged composition over many generations. Here, I used living benthos data from a 47-year time series (1972-2019) alongside bivalve-shell death assemblages sampled in 2008 and 2016 to characterize changes in ecological quality on the Palos Verdes shelf via two index methods: the regionally-developed Benthic Response Index (BRI) and the globally-employed ATZI's Marine Benthic Index (AMBI). Bivalve-generated index scores were positively correlated with scores using the whole fauna, and adjusting bivalve scores from regression coefficients notably improved their agreement of qualitative ecological quality status. Both indices demonstrated that ecological quality had improved from distinctly degraded (1970s-1980s) to recovered (2000s-2010s). Bivalve death assemblages from 2008 indicated that ecological quality then was as high as or higher than that registered by the living bivalve community, and death assemblages from 2016 indicated continued improvement. The discordance between these 21^{st}-century death assemblages with living assemblages from the peak-pollution interval in the mid-20^{th} century suggests that, at least when measured using these biotic indices, death assemblages have a very short memory of past conditions, apparently being overwhelmed by the disturbance-sensitive taxa that have increasingly dominated the shelf in the last few decades.

Records of Biological Invasion. Short-lived biological invasions may leave lasting impacts on ecosystems well after they have concluded, yet the nature of such events is difficult to elucidate in the absence of data from ambient or targeted monitoring efforts. Here, the

ability for surficial death assemblages to recount such invasion events and their ecological legacies was tested using mangrove-dwelling bivalves from Kingston Harbour, Jamaica, where the Asian green mussel *Perna viridis* was introduced ∼20 years ago. While rare in Kingston Harbour today, relative densities of dead P. viridis shells mapped well to historic surveys from early into the invasion and thus help reconstruct spatial variations in invasion intensity. Live-dead discordance of the epifaunal bivalve community further indicated that species have not returned to pre-invasion relative abundance distributions: the economically-important mangrove oyster (*Crassostrea rhizophorae*) has notably declined while the flat tree oyster (*Isognomon alatus*) rose to dominance. Finally, the newly-introduced charru mussel (*Mytella strigata*) in Kingston Harbour has not yet been significantly incorporated into the subfossil record. This case study exemplifies the utility of underexploited sources of geohistorical data for informing the growing problem of human-assisted biological invasion.

CHAPTER 1

LIVING AND DEAD BIVALVES ARE CONGRUENT SURROGATES FOR WHOLE BENTHIC MACROINVERTEBRATE COMMUNITIES IN PUGET SOUND

1.1 Abstract

To integrate paleoecological data with the 'whole fauna' data used in biological monitoring, analyses usually must focus on the subset of taxa that are inherently preservable, for example by virtue of biomineralized hardparts, and those skeletal remains must also be identifiable in fragmentary or otherwise imperfect condition, thus perhaps coarsening analytical resolution to the genus or family level. Here we evaluate the ability of readily preserved bivalves to reflect patterns of compositional variation from the entire infaunal macroinvertebrate fauna as typically sampled by agencies in ocean monitoring, using data from ten long-established subtidal stations in Puget Sound, Washington State. Similarity in compositional variation among these stations was assessed for five taxonomic subsets (the whole fauna, polychaetes, malacostracans, living bivalves, dead bivalves) at four levels of taxonomic resolution (species, genera, families, orders) evaluated under four numerical transformations of the original count data (proportional abundance, square root- and fourth root-transformation, presence-absence). Using the original matrix of species-level proportional abundances of the whole fauna as a benchmark of "compositional variation," we find that living and dead bivalves had nearly identical potential to serve as surrogates of the whole fauna; they were further offset from the whole fauna than was the polychaete subset (which dominates the whole fauna), but were far superior as surrogates than malacostracans. Genus- and family-level data were consistently strong surrogates of species-level data for most taxonomic subsets, and correlations declined for all subsets with increasing severity of data transformation, although this effect lessened for subsets with high community even-

ness. The strong congruence of death assemblages with living bivalves, which are themselves effective surrogates of compositional variation in the whole fauna, is encouraging for using bivalve dead-shell assemblages to complement conventional monitoring data, notwithstanding strong natural environmental gradients with potential to bias shell preservation.

This chapter was originally published as: Kokesh B.S., Burgess D., Partridge V., Weakland S., and Kidwell S.M. Living and dead bivalves are congruent surrogates for whole benthic macroinvertebrate communities in Puget Sound. *Frontiers in Ecology and Evolution*, 10:980753, 2022. doi:10.3389/fevo.2022.980753.

1.2 Introduction

Normalizing the application of paleoecological and other geohistorical data to the direct observations of living systems central to conservation biology is an important, oft-cited objective of conservation paleobiology (Dietl and Flessa, 2011; Kidwell and Tomasovych, 2013). Biological monitoring programs are an important component of those observations and require consistent sampling methods, expert taxonomic training, and rigorous quality assurance procedures, especially where regulated by government agencies (e.g., Dutch et al., 2018). The standards for monitoring benthic communities typically include a wide taxonomic breadth (e.g., macroinvertebrate infauna), species-level taxonomic information, and precise numerical estimates of animal density (e.g., raw counts per taxon standardized to benthic area or sediment volume).

Paleoecological data are typically limited in comparability to such monitoring data in several ways. First, death assemblages – i.e., dead and discarded remains encountered during sampling for living individuals – and fully-buried fossil assemblages are almost always restricted to taxa that possess durable hardparts, such as calcifying mollusks, arthropods, and bryozoans. Under oxygenated waters, soft-bodied clades have low preservation potential, an effect that has rarely been fully quantified (but see the seminal analyses by Schopf,

1978 and Staff et al., 1986). Second, owing to fragmentation and other damage accrued during accumulation in the surface mixed layer, not all dead shell specimens are identifiable to the species level, requiring analytic coarsening to, for example, genus or family level (e.g., Kowalewski et al., 2003; Lloyd et al., 2012; Albano, 2014). Finally, some museum and sedimentary archives might only yield presence-absence (occurrence) data for species, in contrast to the numerical counts generated by bulk sampling, spurring efforts by paleontologists to upgrade data *post-hoc* (e.g., Harnik, 2009; Close et al., 2018). Successful integration of paleoecological data with modern biological data thus requires confidence that assemblages of skeletal remains alone can serve as reliable proxies of the broader fauna targeted for study – usually the entire macrobenthos or at least its infaunal portion – and yield reliable versions of the biological metric of interest (e.g., richness, compositional variation, trophic structure).

Fortunately, the challenges that these three limitations place on the reliability of paleoecological data are directly analogous to well-studied issues of "surrogacy" in the biomonitoring literature (Moreno et al., 2007). This term encompasses an array of techniques aimed to reduce the cost (time, labor) and/or expertise required to process biological samples while still attaining adequate biological insights. That is, referring to our challenges above, acquiring an accurate or sufficient picture when only a subset of clades can be sampled, when species-level resolution might not be possible, and/or when numerical abundance is not necessarily trustworthy. Surrogacy falls broadly into two categories (Table 1.1). 'Sufficiency' refers to coarsening the detail of information collected, either via 'taxonomic sufficiency' *sensu* Ellis (1985; aggregating species-level data into higher taxonomic ranks, functional guilds, etc.) or via 'numerical sufficiency' (simplifying organismal counts into ranked abundances, categorical abundances, or simply presence-absence data). 'Subsetting' on the other hand refers to narrowing the scope of information, for example by focusing on a 'taxonomic subset' (a singular clade, functional guild, or other group of interest) or on a 'numerical subset' of the whole fauna (e.g., the first 100 individuals picked from the sample, as is common in

micropaleontology). Other forms of surrogacy include "cross-taxon" approaches in which one group of taxa serves as a proxy for another, completely independent group (Mellin et al., 2011; Gladstone et al., 2020). A death assemblage, which is not part of the living fauna, could be considered a cross-taxon surrogate that is also a temporally coarser sample than the corresponding living assemblage. Surrogacy methods have been the subject of many tests for macroinvertebrate communities from marine (Warwick, 1988; Ferraro and Cole, 1992; Włodarska-Kowalczuk and Kędra 2007, Bevilacqua et al. 2009), freshwater (Jones, 2008; Mueller et al., 2013; Heino, 2014), and terrestrial settings (Pik et al., 1999; Timms et al., 2013; Souza et al., 2016). Several studies have examined the sensitivity of surrogacy techniques on compositional variation in the fossil record (e.g., Pandolfi, 2001; Forcino et al., 2012; Zuschin et al., 2017), on the faithfulness of dead-shell assemblages to their living counterparts (e.g., Albano et al., 2016), and on the preservation of broader macrofaunal patterns by the preservable subset (Tyler and Kowalewski, 2017).

Here, we evaluate the individual and interacting effects of three surrogacy methods – taxonomic sufficiency, taxonomic subsetting, and numerical sufficiency (Table 1.1) – as well as the effects of data transformation using an infaunal macrobenthic dataset from ten sediment monitoring stations in Puget Sound, Washington State. Spanning 25 classes and 15 phyla, the numerical, species-level data were generated as part of a regional, state-mandated monitoring program of sediments, conducted under strict standards for processing and taxonomic consistency (Dutch et al., 2018). Bivalve death assemblages were rescued from the same sediment grab samples processed for living assemblages, permitting us to directly test the comparability of the preservable subset of the fauna with the whole fauna, including the live-dead agreement of the bivalve subset of that fauna. Long-term biomonitoring has demonstrated that these ten subtidal sites have been remarkably stable in community composition on this regional scale: scatter in ordination space produced by temporal variation over the last three decades is small relative to the separation among stations, despite other

evidence of overall deteriorating conditions across Puget Sound (Partridge et al., 2018). This tension of compositional stability, and the possibility of using death assemblages for insights into community states before the onset of monitoring, motivated this collaboration: the test of shelly fauna as surrogates of the whole fauna is an important first step both in assessing the reliability of fossil data for local use and developing a protocol for wider application.

1.3 Methods

1.3.1 Study area

Puget Sound is a fjordic estuarine system in Washington State that, along with the Strait of Georgia and Strait of Juan de Fuca, makes up the Salish Sea (Figure 1.1), and was formed by the southernmost lobe of the Cordilleran Ice Sheet during the Fraser and Vashon glaciations (∼20-15 ka; Booth, 1994). Puget Sound is the second-largest estuary in the United States by area (2,600 km2) after Chesapeake Bay (11,600 km^2), but has considerably greater mean and maximum water depths (140 m and 280 m, respectively) than Chesapeake Bay (7 m and 53 m) and accommodates 250% more water by volume. Macrotidal flushing from the Sound through the Strait of Juan de Fuca to the Pacific Ocean facilitates the import of cold bottom-waters that circulate into major basins of the Sound, with the result that sediment-dwelling macrobenthos encounter fully saline, oceanic waters notwithstanding freshwater-influenced surface lenses (Moore et al., 2008). This flushing is complicated, however, by a series of narrow, silled chokepoints between basins that slow the circulation of water considerably towards the terminal ends of the Sound (Khangaonkar et al., 2011; but see MacCready et al., 2021). The numerous coastal inlets and deep basins within the greater Puget Sound area (US waters of the Salish Sea, henceforth 'Puget Sound' for brevity) thus provide a complex array of intertidal, shallow subtidal, and deep subtidal habitats in a hydrographic estuary, with diverse sediment facies and naturally variable hydrologic conditions (e.g., dissolved oxygen,

salinity). The Sound also presents a complex pattern of development in adjacent watersheds, which range from rural to suburban, industrial (historically ore, forestry, military), and densely urbanized cities such as Seattle and Tacoma (Figure 1.1).

Benthic macroinvertebrate communities in Puget Sound have been relatively well-documented since the 1960s by a combination academic (Lie and Evans, 1973), federal (Nichols, 2003), and, over the last 30 years, state agencies (Dutch et al., 2018). Since 1989, ten 'sentinel' stations located across Puget Sound have been monitored by the Marine Sediment Monitoring Team at the Washington State Department of Ecology (ECY), focusing on benthic communities, sedimentology, and chemical contaminants as part of the Puget Sound Ecosystem Monitoring Program (PSEMP). These ten sentinel stations were selected from a larger set of stations sampled until 1995 because they represent distinct habitats from which distinct benthic communities were observed. PSEMP data reveal long-term declines in benthic condition based on a 'sediment quality triad index' that uses community composition (whole fauna), sediment toxicity (using echinoid fertility and amphipod mortality tests), and a suite of chemical contaminants (Long and Chapman, 1985). However, causal mechanisms for this decline have been difficult to identify (Partridge et al., 2018; Weakland et al., 2018). The PSEMP protocol has thus recently been revised to assess a broader suite of variables including macroinvertebrate biomass, nutrients, and stable isotopes to investigate additional possible stressors (Dutch et al., 2018). The overall decline in Puget Sound benthic quality detected by PSEMP since 1989 involved increased toxicity despite no change in contaminant levels, and significant declines in richness and total invertebrate abundance in several areas within the large network of sites (Partridge et al., 2018). However, the annually-sampled set of ten sentinel sites used here exhibited relative stability over this period, despite strong fluctuations in different taxa at different sites: habitat-level differences in sediment grain size and water depth are the strongest correlates of community composition, with date of sampling a secondary and much smaller effect (Figure 7 in Partridge et al., 2018).

1.3.2 Data preparation

The dataset of living assemblage composition used here comes from ECY samples acquired in the years 2017 and 2018, aggregated to achieve adequate sample sizes, at each of ten, long-established monitoring stations in Puget Sound (Figure 1.1), using the entire living macroinvertebrate community (henceforth 'whole fauna'). Sampling followed the Puget Sound Estuary Protocol (PSEP, 1987), which is standard among many agencies along the US Pacific Coast. Samples were collected using a 0.1-m^2 van Veen grab with a minimum sediment penetration of 5 cm (Puget Sound Estuary Program, 1987). Sediments were sieved through 1-mm mesh on deck with seawater, fixed in 10% formalin, and then exchanged into 70% ethanol within 3-7 days. ECY taxonomists sorted and identified all invertebrates to the species level, when possible, with reference to a library of voucher specimens; coordination with taxonomists at other agencies in the Pacific US (SCAMIT, 2021) has also contributed to taxonomic consistency.

Dead bivalves were sorted from the dried sediment residue of two years of benthic samples at the same sites collected in the years 2018 and 2019: the death assemblage had not been saved from the 2017 samples and the living whole fauna from 2019 was not yet available. This sediment residue is typically discarded by agencies after living macroinvertebrates have been removed. All bivalve shells with an identifiable hinge were counted as individuals, whether the valve itself was whole or fragmented, and were identified taxonomically to the finest resolution possible. Each fully disarticulated valve or hinge fragment was considered as coming from a different living individual, following the reasoning of Gilinsky and Bennington (1994) and standard protocol for live-dead analysis of mollusks of Kidwell (2009): no correction for the number of skeletal units per living individual (e.g., Kowalewski and Hoffmeister, 2003) was needed because only bivalves were considered. However, to avoid conflating the population-density data for living bivalves and time-averaged abundance data for dead bivalves, we avoided direct comparisons of raw counts and instead used proportional

abundances for all analyses unless otherwise indicated.

From the original whole fauna dataset, new species-level abundance tables were extracted analytically for polychaetes, malacostracans, and bivalves. A fifth table was produced using species-level abundance data from the rescued dead-shell assemblages. Species from each of these five tables were then coarsened to create abundance tables for genera, families, and orders. Each abundance table was then subjected to four different data transformations that successively reduce the influence of dominant taxa and increase the influence of rare taxa: non-transformed proportional abundances, square-root transformed, fourth-root transformed (the preferred transformation adopted by ECY and widely used in surrogacy studies; Partridge et al., 2018), and binary presence-absence. While the first three of these transformations preserve the abundance structure of the data, presence-absence represents a true test of numerical sufficiency as it does not rely on knowledge of the underlying count data. We note that other transformations commonly applied to ecological data, such as log and arcsine, could also be tested. Our focus on the series x^1, $x^{0.5}$, $x^{0.25}$, and x^0 instead allows us to measure the effects of progressively severe transformations – increasing deviation from the original abundance structure – using the same mathematical operator. A log transformation should moreover not be judged a priori as more or less severe than an n-root transformation (although log(x+1) has been argued to be equivalent to $x^{0.25}$; Clarke and Warwick, 2001), whereas we can be confident that fourth-root transformed data will always be less correlated to non-transformed data than will be square-root transformed data. Additional transformations that would qualify as methods of numerical sufficiency such as semi-quantitative ordinal abundances(e.g., common, rare) were not assessed because the number of ordinal categories and their abundance ranges are largely at the discretion of the investigator.

Distance matrices for each of these surrogate abundance tables were generated using Bray-Curtis dissimilarities (Jaccard in the case of presence-absence) among the ten monitoring stations. This work produced 80 distance matrices (5 subsets $\times$ 4 resolutions $\times$ 4

transformations), with 79 of these serving as 'surrogates' of the original, benchmark matrix (proportional abundances at the species level for the whole fauna).

1.3.3 Data analysis

Richness (the raw count of taxa), abundance (raw number of individuals), and evenness (abundance distribution among taxa) for each surrogate dataset were calculated both at the station scale (i.e., at each of the 10 stations) and at the regional scale, that is, after aggregating station-level abundances together. Richness values were converted into the ratio of higher taxa over the number of species (T/S) for genera (G/S), families (F/S), and orders (O/S). T/S values range from >0 to 1, with 1 indicating that all members of the higher taxon are monospecific. Abundance values were converted into the ratio of the abundance of individuals in the subset over the abundance in the whole fauna (N/W). N/W values range from >0 to 1; a value of 0.5 would indicate that the subset represents 50% of all individuals in the whole fauna. We expressed evenness as the probability of interspecific encounter (PIE; Hurlbert, 1971), which is relatively insensitive to sample size and richness >∼5 (Olszewski, 2004). PIE values have a potential range from 0 to 1, with 1 indicating a perfectly even community (same number of individuals per taxon) where individuals of any species are equally likely to be randomly sampled. T/S, N/W, and PIE were compared among surrogate datasets using the regional-scale values and the medians and interquartile ranges (IQR) of station-scale values.

Spearman rank correlation coefficients (ρ) between all pairwise combinations of Bray-Curtis distance matrices were calculated and tested for significance using Mantel tests with 999 permutations (Mantel, 1967). Coefficients can range from -1 to +1. A pairwise correlation of +1 would indicate that the Bray-Curtis distances among stations for two surrogate datasets are in perfect rank order: put another way, the two surrogates would produce identical non-metric multidimensional scaling ordinations (NMDS). We assessed the effects of

each surrogacy method on the perception of compositional variation among sites (see Table 1.1) by evaluating ρ values varied by treatments among the full set of pairwise comparisons. First, taxonomic sufficiency was evaluated using the ρ values generated by comparing matrices of species-level data with those for higher taxa while holding both the subset and data transformation level constant. Second, taxonomic subsetting was evaluated by using the ρ values generated by comparing the whole fauna and each surrogate subset while holding both the taxonomic resolution and data transformation level constant. Third, the effects of data transformations (and numerical sufficiency in the case of presence-absence data) were evaluated using the ρ values generated by comparing matrices based on proportional abundances with those based on more severe data transformations while holding both taxonomic resolution and subset constant. Finally, the composite effect of all three surrogacy methods was evaluated using ρ values generated by comparing the original dataset (species-level whole fauna using proportional abundances) to all surrogate matrices.

In a procedure known as a second-stage analysis introduced by Somerfield and Clarke (1995), pairwise ρ values were themselves compiled into distance matrices by rescaling ρ such that $D_\rho = (1 - \rho) \,/\, 2$, where D_ρ is the distance between two surrogates (see Figure 1.2 for schematic diagram). A coefficient of perfect correlation ($\rho = +1$) thus corresponds to a $D_\rho = 0$. The new distance matrices were then used to generate 'second-stage' NMDS. Each point in this ordination represents an entire surrogate's distance matrix, not the positions of individual samples that are plotted by first-stage, 'conventional' ordinations. The more closely that two points plot to each other, the higher their pairwise ρ value relative to that with other surrogates. Second-stage NMDS was performed separately for each of the five taxonomic subsets (16 matrices each subset, i.e., 4 resolutions $\times$ 4 data transformations) and then re-run to combine all 80 surrogates simultaneously. Differences between the multivariate means of all surrogates grouped by taxonomic resolution, subset, and data transformation were tested using 1-way analysis of similarity (ANOSIM; Clarke, 1993) based on the ranked

distances of the matrix used to generate the combined second-stage NMDS. To test whether patterns observed from the combined second-stage NMDS were consistent across the 30-year time series of available community data, we performed analyses for each 2-year aggregation of the living community and calculated the correlation between resulting second-stage distance matrices to that of the most recent living community (2017-2018).

All analyses were conducted in the statistical environment R version 4.2.0 (R Core Team, 2023).

1.4 Results

1.4.1 Taxonomic structure of the macrobenthic fauna

The whole macrobenthic fauna sampled in the two annual samples of 2017 and 2018 comprised 22,713 individual organisms drawn from 319 species, 240 genera, 146 families, 74 orders, 25 classes, and 15 phyla (Table 1.2). Polychaetes were the most abundant (4,817 individuals, 43% of the whole fauna) and diverse class (157 species, 49% of the whole fauna). Malacostracans were second in diversity (61 species, 19% of the whole fauna), but third in abundance (433 individuals, 4% of the whole fauna). Bivalves were third in diversity (33 species, 10% of the whole fauna), but second in abundance (4,367 individuals, 39% of the whole fauna). Dead bivalves acquired from annual samples in 2018 and 2019 comprised 5,337 individuals and 43 species.

T/S ratios predictably and significantly declined from genera to orders based on visual separation of IQRs (top row, Figure 1.3). Trends in T/S for the whole fauna and polychaetes steadily declined to ratios of 0.23 and 0.12, respectively. Malacostracans had a notable decline between families (0.59) and orders (0.11). Living and dead bivalves had similar station- and regional-scale patterns and declined less strongly than did other subsets and the whole fauna. Polychaetes, living bivalves, and dead bivalves had similar regional-scale

N/W ratios (middle row, Figure 1.3). Note that N/W ratios do not change at different taxonomic resolutions, because the same number of individuals are used regardless of the resolution assessed. PIE values at the station- and regional-scale were higher and less variable for the whole fauna and polychaetes compared to malacostracans and bivalves and exhibited scant differences among taxonomic resolutions (bottom row, Figure 1.3). Regional-scale PIE values for malacostracans were high from the species to family levels (>0.9), but declined for orders (0.8). Living and dead bivalves had the lowest PIE values at the regional-scale (<0.8), and also exhibited no differences among taxonomic resolutions.

1.4.2 Effects of surrogacy techniques

Focusing on the effect of taxonomic sufficiency (see Table 1.1): Spearman ρ values between a distance matrix based on species-level data and those based on higher taxonomic resolutions varied by subset but were minimally affected by data transformation (top row, Figure 1.4). All surrogate matrices of the whole-fauna had high ρ values (>0.8) with their species-level version. For polychaetes, genera had strong correlation to species, but ρ values dropped for families and orders. Correlations also declined strongly with taxonomic resolution for malacostracans, with ρ values for orders ranging from 0.25 to 0.6, depending on the severity of data transformation. Living bivalves had strong correlations up to the family level, but then declined slightly for orders. Finally, dead bivalves had strong correlations with their species-level version at all taxonomic resolutions, although ρ values notably declined at the genus level using proportional abundance data. All correlations were significantly positive based on Mantel permutation tests ($p < 0.05$).

Concerning the effect of taxonomic subsetting, Spearman ρ values between the whole fauna and the four taxonomic subsets demonstrated that polychaetes were the strongest surrogate for most levels of taxonomic resolution and data transformation (second row, Figure 1.4). Living and dead bivalves exhibited moderate ρ values to the whole fauna, and

malacostracans were notably weak surrogates for all resolutions and data transformations (all $\rho < 0.4$). Most surrogate matrices for malacostracans were not significantly correlated to their whole-fauna counterparts.

Concerning the effect of numerical sufficiency, Spearman ρ values between a matrix of proportional abundance data and all other transformations of that same subset and resolution demonstrated consistent weakening of correlation with increased severity of data transformation (third row, Figure 1.4). Generally, square-root transformed data were strongly correlated to proportional abundance data, whereas fourth-root transformed data and presence-absence data each decreased correlations detectably. Correlations using presence-absence data for living and dead bivalves were particularly low, with several species- through order-level comparisons losing significance.

For the composite effect of all three surrogacy methods, Spearman ρ values for comparisons with the original dataset, all three methods had detectable effects (bottom row, Figure 1.4). Surrogates using polychaetes had strong to moderate ρ values for all resolutions and data transformations. Surrogates using malacostracans exhibited the weakest correlations, most of which failing to demonstrate significance. Surrogates using living and dead bivalves were moderately correlated with the original dataset except when using presence-absence data.

1.4.3 Comprehensive patterns of inter-matrix correlations

For each taxonomic subset, second-stage NMDS plots of inter-matrix correlations – where points separate according to the distinctness of surrogate datasets, not samples within a surrogate dataset – all produced ordinations where the axes were reflected by the effects of taxonomic and numerical sufficiency (top row, Figure 1.5). The first NMDS axis for the whole fauna, living bivalves, and dead bivalves were ordered by the severity of data transformation (proportional abundances to presence-absence) from left to right. The second NMDS axis

for these three sets generally defined a gradient of taxonomic resolutions, with species and genera at the negative end of the axis (bottom of plots) and families and orders at the positive end (top). Polychaetes and malacostracans exhibited a "fanning-out" pattern such that species-level matrices were positioned in the lower right quadrant of each plot and higher taxa of different data transformations progressively spread out towards the left.

For the combined second-stage NMDS – where all points (surrogate datasets) are displayed in a shared ordination space – the first NMDS axis was defined by clusters of taxonomic sets (bottom row, Figure 1.5). From left to right, living bivalves, dead bivalves, polychaetes, and the whole fauna were only slightly separated from each other at the negative end of the axis, whereas malacostracans plotted on the positive end of the axis, far from all other subsets. The second NMDS axis generally defined a gradient of data transformation severity, with matrices based on proportional abundances plotting at the negative end and those based on presence-absence plotting on the positive end. This vertical gradient was apparent for all sets except malacostracans: their matrices differentiated by taxonomic resolution (i.e., same set and data transformation) formed tight clusters with no apparent directionality. ANOSIM tests of inter-matrix differences in correlations (Table 1.4) corroborated the visual patterns from second-stage NMDS. Differences among matrices were significant based on taxonomic subset ($R = 0.74$, $p < 0.001$) and less strongly so for data transformation ($R = 0.08$, $p = 0.003$). Differences among matrices based on taxonomic resolution were not significant (Table 1.4).

Finally, rerunning the second-stage NMDS analyses for all two-year sets of the living assemblages from the full 30-year time series produced the same overarching pattern as described in detail above for the 2017-2018 set of living assemblages. Polychaetes were always the strongest surrogate subset of the whole fauna, followed by bivalves, and trailed by poorly-performing malacostracans. Importantly, in each of these replicate tests, the distance matrices produced by bivalve death assemblages collected in 2018-2019 were consistently

congruent with those of living bivalves. The consistency of results across the 30-year time series is summarized succinctly by high correlation coefficients ($\rho \geq 0.8$) between second-stage distance matrices produced using older sets of living assemblages and that produced using the 2017-2018 set (Figure 1.6).

1.5 Discussion

The aim of this study was to assess the ability of three surrogacy methods – all commonly adopted to streamline sample processing – to predict the original inter-station compositional distances exhibited by the whole fauna: are the compositional distances among stations observed using species-level data for the whole fauna retained when data are coarsened to genus or higher resolution, and/or transformed toward presence-absence, and/or subsetted to focus on a single class? In addition, how closely does the local dead-shell assemblage (and its surrogates) mirror the counterpart subset of the whole fauna, relevant to using this usually-discarded part of benthic samples for insights into ecological history? Methods of surrogacy were evaluated for both their individual effects and their effects in various combinations, and they were applied to the class Bivalvia using both living and dead-shell-based data. Importantly, our analyses assess the ability of surrogates to preserve compositional distances among stations over a regional scale: we do not assess the ability of surrogates to detect a known spatial or environmental gradient (i.e., contrary to Tyler and Kowalewski, 2017, and Kokesh et al., 2022) nor their power when formulated as benthic quality metrics (e.g., Tweedley et al., 2014; Dietl et al., 2016; Pruden et al., 2021).

Our analyses could be improved by including additional monitoring stations despite the wide range of depths and substrates the dataset encompasses within Puget Sound (Table 1.2). However, counterpart analyses of the ability of surrogates to capture regional patterns are definitely needed in other study systems, the better to assess the power of surrogates on local scales, including the generality of the results here. Our discussion focuses on directions

needing additional research and some ways to accelerate that progress, such as focusing analyses on gradients in clearly important biological factors, such as the number of species per higher taxon, relative abundances of subset clades, and community evenness.

1.5.1 Taxonomic sufficiency: Coarser resolutions are effective surrogates

Variation in compositional distances among stations was reliably preserved despite the coarsening of taxonomic resolution using most taxonomic subsets – i.e., all correlations were significant – even though ρ values did decline with increased coarsening (first row, Figure 1.4). These results corroborate the findings of numerous earlier studies that inter-matrix correlations do decline with increased taxonomic coarsening (Olsgard and Somerfield, 2000; Włodarska-Kowalczuk and Kędra, 2007; Bertasi et al., 2009; Bevilacqua et al., 2009), but that family-level identification generally provides a suitable approach to taxonomic surrogacy for the whole fauna (Jones, 2008; Heino, 2010; Pitacco et al., 2019). We found that family-level data were comparable to genus-level data for both living and dead bivalves, even though it was weaker than genus-level data for polychaetes and malacostracans. The observed variation among subsets in their robustness to taxonomic coarsening suggests that one should test their retention of power prior to enacting this surrogacy method into practice. In addition, taxonomic sufficiency is a relatively conservative surrogacy method compared to taxonomic subsetting (see next section), as demonstrated here via second-stage NMDS (Figure 1.5) and discovered earlier by Mellin et al. (2011) using meta-analysis.

The declining power of surrogates with taxonomic coarsening observed here is likely related to the packing of species into higher taxa (e.g., Heino and Soininen, 2007; Bevilacqua et al., 2012; de Oliveira et al., 2020). We found that the decline of ρ values with coarsening resolution (top row, Figure 1.4) were commensurate with declines in the T/S ratio (top row, Figure 1.3). Others have postulated that a T/S ratio ≥ 0.4 (also reported as a S/T ratio ≤ 2.5) signals that the higher taxon will be able to sufficiently mirror patterns among samples

detected using species-level data (Timms et al., 2013). That approximate value has also been identified as a threshold in other systems (Albano et al., 2016; Kokesh et al., 2022). This study thus adds support for the growing consensus that there exists a minimum T/S ratio above which higher taxa are likely to be effective surrogates for species-level patterns, providing a rapid means of predicting success, although this threshold might vary among biological systems (Heino, 2014).

Relevant to the paleoecological potential of taxonomic sufficiency, our finding of particularly strong correlations of species to higher taxa of dead bivalves at multiple resolutions supports the common practice of evaluating fossil assemblages at the resolution of genera or families: such coarsening produces compositional patterns similar to those of species-level information, as also found by others using exclusively fossil data (Forcino et al., 2012; Zuschin et al., 2017). The persistence of live-dead patterns in Puget Sound despite this small magnitude of taxonomic coarsening on live-dead patterns has also been recognized elsewhere (e.g., Albano et al., 2016). We suspect, however, that the power of taxonomic sufficiency will vary over space and time in both modern and ancient faunas, including a likely latitudinal gradient. For example, S/G ratios of living bivalves are particularly high in the tropics and low in boreal settings (Krug et al., 2008), both regions where relatively few sufficiency tests have been conducted compared to mid-latitude settings.

1.5.2 Taxonomic subsetting: Well-represented subsets are effective surrogates

Subset clades exhibited disparate strength as surrogates for the whole fauna, but polychaetes were the most powerful (second row, Figure 1.4), as has been found in numerous tests of class-level subsets in marine benthos (Olsgard et al., 2003; Włodarska-Kowalczuk and Kędra, 2007; Kokesh et al., 2022). This high fidelity of polychaetes is usually attributed to their comprising the majority – sometimes the overwhelming majority – of individuals and species richness in soft-sedimentary seafloors. However, polychaetes were closely followed in

Puget Sound by bivalves, both alive and dead, by total abundance and were only distantly trailed by malacostracans, despite that group being second-ranked in richness in the Puget Sound regional dataset (second row, Figure 1.4). In general, taxonomic subsetting has been generally found to be less effective, reliable, and/or consistent as a method of surrogacy than has taxonomic sufficiency (coarsening), based on other studies using the same clades as examined here (e.g., Włodarska-Kowalczuk and Kędra, 2007; Bevilacqua et al., 2009; Mellin et al., 2011; Gladstone et al., 2020). Nonetheless, given the observed variation among subsets in their robustness to taxonomic coarsening, we would suggest that their ability to preserve patterns be assessed prior to putting this (or any other) surrogacy method into practice. Thus, although the results here are very positive for bivalves in particular, their reliability should be tested locally prior to formalizing surrogacy in a protocol. Such tests could repeat the same methods described here but use a limited number of stations or a single historic survey to produce a benchmark whole-fauna matrix by which to test for the reliability of subsets.

Additional research is also needed to identify the biological properties of taxonomic subsets that convey reliable surrogacy power: these determinants are underexplored compared to those that allow a subset to retain power despite coarser resolution (previous section). One promising attribute is the proportional abundance of the subset within the whole fauna (e.g., Dietl et al., 2016). For example, we found that taxonomic subsets declined in ρ values (second row, Figure 1.4) congruent with declines in their median station-scale abundance within the whole fauna (N/W; middle row, Figure 1.3): polychaetes (highest N/W) were a more powerful surrogate subset of the whole fauna than were either living or dead bivalves, which were both more reliable than malacostracans (lowest N/W). Malacostracans are represented by more species than bivalves in this dataset, but the potential leverage of that diversity apparently cannot compensate for their low station- and regional-scale abundances. Thus, the ability of taxonomic subsets to serve as effective surrogates appears to

be, at least partially, a function of their relative abundance rather than their absolute or proportional richness in the whole fauna. It would be worthwhile to explore whether some minimum threshold N/W exists that determines the utility of taxonomic subsets as surrogates, even though this threshold might vary among biological systems (especially as a function of community evenness or S/G ratios).

From this single test in Puget Sound, it is difficult to determine the importance of the (quite large) sample size of dead bivalves (N = 5,337) on their performance as a surrogate. Dead bivalves are not technically a subset of the living fauna but rather a cross-assemblage surrogate of much coarser temporal scale. Moreover, their raw abundance might be genuinely biased upward relative to the raw abundance of living bivalves due to our counting disarticulated valves, although that effect on outcome is minimized by our using proportional abundance data). Nevertheless, the high compositional congruence between living and dead bivalves in Puget Sound, as judged by their overlapping clusters in ordination space (bottom row, Figure 1.5), indicates that bivalve death assemblages have a power similar to that of living bivalve subsets for recognizing compositional variation in the whole fauna, that is for serving as surrogates of the whole fauna (see 1.5.4 below).

1.5.3 Numerical sufficiency and data transformation: The importance of abundance data

We found that the severity of data transformation had a much larger effect on the strength of different surrogates than did the coarsening of taxonomic resolution (third row, Figure 1.4). Increasing the strength of data transformation resulted in sequential declines in correlation with the original proportional abundances regardless of taxonomic (sub)set or resolution, and downgrading abundance to presence-absence data had a particularly strong impact for both living and dead bivalves, resulting in correlations that lacked statistical significance. Although matrices based on different data transformations were notably separated in second-

stage NMDS space (Figure 1.5), these effects were small compared to differences created by taxonomic subsetting (Table 1.4).

Ecologists commonly transform data from their original, raw counts in order to better meet the assumptions of univariate tests (usually an assumption of normality; St-Pierre et al., 2018). All such transformations (other than proportional abundance) diminish the influence of dominant taxa on compositional patterns and specifically amplify the role of rare taxa – effectively, this operation artificially inflates the evenness of a community. Like the predictive power of T/S for taxonomic sufficiency discussed above, the evenness of the raw count or proportional abundance data matrix of a surrogate will determine the extent to which data transformation will alter its composition. For example, the relatively low PIE values (low evenness) observed for bivalve living and (especially) death assemblages here (bottom row, Figure 1.3) indicate that data transformation should have a particularly large effect, leveling the initially uneven effect of dominant and rare species on composition until, with the extreme of presence-absence data, all species have equal individuals (PIE = 1). For bivalves in Figure 1.4 (right two columns), data transformation in fact does have a strong effect, accompanied by large decreases in ρ. In contrast, the whole fauna and polychaetes both had high PIE values relative to bivalves (and malacostracans), and their correlation (resemblance) to the original fauna were much less affected by data transformation, although not immune to it (Figure 1.4). The particularly large decline in ρ values for bivalves (and for order-level malacostracans; Figure 1.3) appears to correspond to an outsized response to data transformation by assemblages having relatively low evenness (PIE $< \sim$0.6). Thus, PIE values for a surrogate's raw count data can likely serve as an easily-acquired first-order predictor of the magnitude of effect that data transformation will have on community patterns.

Although square-root and fourth-root transformations are commonly employed for biological data analyses (e.g., Thorne et al., 1999; Clarke and Warwick, 2001; Dethier and Schoch,

2006; Domínguez-Castanedo et al., 2007; Vijapure and Sukumaran, 2019), these particular transformations do not themselves constitute true surrogates: they still require that someone generated the underlying original abundance data. In contrast, ranked-abundance (stepwise integers), ordinal abundance data (e.g., "rare", "common", "very common", etc.), and presence-absence data do represent methods that reduce effort during the data collection, and thus are all potential surrogacy techniques (e.g., Carneiro et al., 2010; Landeiro et al., 2012). It should be noted that using numerical sufficiency as a surrogacy method comes at the cost of reducing the range of possible analyses. For example, richness and binary distance metrics like Jaccard similarities can be assessed accurately, but abundance-based metrics such as rarefaction, similarity percentages (SIMPER; Clarke, 1993), and evenness and dominance metrics become inviable. The objectives of a study or monitoring program, the pre-determined scope of analytical methods, and the results from preliminary tests of consistency among data transformations thus should all be considered in order to determine whether the implementation of numerical sufficiency is acceptable.

1.5.4 Bivalve live-dead agreement and taphonomic caveats

Compositional variation in bivalve death assemblages among the ten Puget Sound stations, where the whole fauna is documented to have been relatively stable since 1989 (Partridge et al., 2018), was congruent with that of living bivalves based on all taxonomic resolutions and data transformations assessed in this study (Figure 1.5). The ability of dead-shell assemblages to mirror regional variation of bivalve living assemblages in Puget Sound is thus very high, and moreover, is robust to taxonomic coarsening and to multiple levels of data transformation. Bivalve death assemblages have been shown to preserve spatial and environmental gradients in bivalve living assemblages in other systems of relatively pristine or otherwise steady-state conditions (e.g., Warme, 1969; Tomašových and Kidwell, 2009; Martinelli et al., 2016; Hyman et al., 2019). Conversely, poor live-dead agreement in subtidal

settings, commonly assumed to reflect postmortem information loss, is now appreciated as being strongly correlated with a recent history of anthropogenic disturbance sufficient to have offset the composition of the living assemblage from its long-term baseline state, which the surficial time-averaged death assemblage retains a memory of (Kidwell, 2007). Stressors with such power include eutrophication (Kidwell, 2007; Korpanty and Kelley, 2014; Leshno et al., 2015; Gilad et al., 2018; Tweitmann and Dietl, 2018), biological invasion (Yanes, 2012; Chiba and Sato, 2013; Steger et al., 2022), bottom trawling (Kidwell, 2009), various combinations of stressors (Haselmair et al., 2021), and warming-associated environmental changes (Powell et al., 2017; Meadows et al., 2019). We note that while live-dead congruence as found here (i.e., living and dead bivalves producing similarly-structured distance matrices of among-station variation) is not the same as live-dead agreement as conventionally determined (i.e., the living and dead bivalves having similar faunal composition), it nevertheless supports the observational data for a stable community organization over the last 30 years at this regional scale (Partridge et al., 2018). Indeed, the consistent results for all two-year sets of the community time series and high correlation to the 2017-2018 set (Figure 1.6) suggests that patterns of taxonomic surrogacy and live-dead congruence among bivalves have both been stable attributes of the Puget Sound benthos for at least the last 30 years.

Based on annual monitoring, overall benthic conditions in Puget Sound are known to have deteriorated over the last 30 years, especially in terminal inlets with poor water exchange and especially using measures of sediment toxicity (Partridge et al., 2018; Weakland et al., 2018). However, changes in benthic communities at the ten sentinel sites used here have been modest and non-directional using compositional measures: most variation is among sites as a function of habitat type, with only a small secondary effect of the date of sampling (Partridge et al, 2018). That said, the ten sites have each exhibited variability including some strong fluctuations in abundance over the 1989-2015 interval: some sites have shown an increase (or a decrease) in total abundance or richness, and some have experienced an increase (or

decrease) in a particular taxon (but usually different taxa at different sites) and none of these faunal changes are correlated strongly with tracked contaminants (Partridge et al., 2018). This variability and the uncertain causes stimulated a change in ECY's monitoring program to include additional environmental measures and increase the number of sites sampled annually (Dutch et al., 2018). It was also a motivation for our collaboration, bringing death assemblages into the evaluation, and thus for this initial investigation of surrogacy.

In light of this known but highly complex and spatially disparate history of community age at these ten sites, which have nonetheless remained compositionally distinct, the high live-dead congruence observed among bivalves suggests that the death assemblages have experienced little net postmortem bias from processes such as lateral transportation and differential preservation of shells. It also suggests that the time averaging of skeletal materials has been either (1) very short, i.e., encompassing too few generations of shell input to detect any past shifts in community composition that might have occurred before the onset of monitoring, or (2) fairly long, but compositional changes within that interval of death assemblage accumulation did not exceed the range of variation observed within the past 30+ years (i.e., the distance matrices of community composition have been stable for more than 30 years).

We do not yet have shell-age data to test these alternatives directly. Such data on the window of time averaging (that is, the age of the oldest shells) and the temporal resolution (IQR of the frequency distribution of cohorts) of a death assemblage is typically quantified by radiocarbon or other direct age-dating of shells (for review, see Kidwell, 2013). The per-sample cost and subsample size have declined over the last decade, but shell-age analysis still requires funds and collaborations outside those generally on hand. It thus is still undertaken only to address specific issues relevant to management that emerge from an initial live-dead analysis, such as the date of disappearance of an abundant but dead-only taxon (e.g., Kowalewski et al., 2000; Albano et al., 2016; Tomašových and Kidwell, 2017; LeClaire

et al., 2022) and the persistence of an important habitat (e.g., Waldbusser et al., 2013; Casebolt and Kowalewski, 2018; Hyman et al., 2019; Haselmair et al., 2021). The precise scale of time averaging is not critical to the question here, nor has local shell-age data accompanied surrogacy tests by others using live and dead bivalves (e.g., Schopf, 1978; Tyler and Kowalewski, 2017), including surrogacy tests that suppose that dead-shells should serve as good guides to gradients because the living mollusk subset does (Dietl et al., 2016; Kokesh et al., 2022). Age data is in fact lacking in many valuable and almost certainly correct inferences of benthic change based on live-dead discordance (e.g., Tweitmann and Dietl, 2018). Thus, for an initial analysis of the local value of death assemblages, it can suffice to appreciate that dead shells reflect the "past" to some degree, notwithstanding uncertainties in the magnitude of time averaging, contrasting with the "now" represented by living assemblages. This qualitative strategy of analogizing live/dead (or live/Quaternary fossil, or core-top/core-bottom) with now/then underlies much of conservation paleobiologic research and achievements, allowing us to acquire first-order insights before fully knowing the geologic age and age-structure of the local shelly record.

Shell-age analysis is only now getting underway for Puget Sound bivalve assemblages. However, geological inference suggests that time averaging might be quite limited in cold-temperate Puget Sound, perhaps reflecting shell input largely from only the last few decades – that is, encompassing the duration of biomonitoring but perhaps not much more. One argument for such short time averaging is that the ratio of dead-to-live bivalve abundances is only 1.2 to 1 (Table 1.3), which is much lower than the global median of 8:1 for tropical and warm-temperate settings (Kidwell, 2013). This low ratio suggests relatively little accrual of past generations in the surface mixed layer, which can have several non-mutually exclusive explanations. One is that postmortem rates of aragonitic shell disintegration are high, which we would suspect for relatively cold waters like those of Puget Sound. A second explanation for low dead-to-live ratios can be a high rate of sedimentation, which removes shells from the

surface mixed layer via burial. Sedimentation rates in Puget Sound are, in fact, relatively high at ∼0.5-2 cm/yr based on 210Pb dating of sedimentary cores (Brandenberger et al., 2011).

Data from other settings with high sedimentation provide the best basis for predicting the scale of time averaging of Puget Sound bivalve death assemblages, pending direct dating. For example, in subtropical siliciclastics of Texas coastal lagoons, with a sedimentation rate of ∼0.3 cm/yr, nearly all dead bivalve shells from the upper 0-10 cm of sediment were dated as younger than ∼10 years (Olszewski and Kaufman, 2015). In tropical lagoonal deposits of the Great Barrier Reef with a sedimentation rate of ∼0.6 cm/yr, the 95% confidence interval of median shell ages from the uppermost 15 cm of sediment ranged from 2 to 167 years with a maximum shell age of ∼200 years old (Kosnik et al., 2015). Given that these are much warmer-water settings with comparable sedimentation rates, these values probably define the upper limit for time-averaging expected in the cold-temperate Puget Sound. Our best expectation for Puget Sound is thus IQRs on the order of a few decades, with few old shells extending more than ∼100 years old.

1.5.5 Implications for conservation biology and paleoecology

Given that variation in living bivalves (a true subset of the whole fauna) demonstrated strong correlation to the whole fauna, it follows that bivalve death assemblages, which map so closely onto living bivalves, possess a similar capacity to that of living bivalves to approximate the among-station compositional distances of the whole living macrobenthic fauna in this system. Polychaetes are the strongest surrogate subset here (as they also are in many other systems, see references above), but they and species-rich malacostracans lack the preservation potential of bivalves under normoxic conditions and thus have little practical paleoecological potential. Albano et al. (2016) found species-level patterns of live-dead agreement to be consistent up to resolutions of families, corroborated further with regard to

the effects of data transformation in Miocene assemblages by Zuschin et al. (2017). Using a different methodological approach, namely quantifying compositional variability among environmentally-categorized groups of stations, Tyler and Kowalewski (2017) found that molluscan death assemblages preserve the extent of variation among environments as reflected by the whole living fauna. This study introduces, for the first time, methods already in wide use by applied ecologists for examining surrogacy (particularly second-stage NMDS) and expands to include both living and death assemblages. Our results corroborate those of previous fossil-focused investigations.

Our finding of the reliability of bivalve death assemblages supports their use for retrospectively evaluating human impacts on ecological systems. Conservation biologists typically have to rely on long-term monitoring data to quantify ecological responses to historic human impacts, watching changes (decline, recovery) play out in real time (e.g., Borja et al., 2006; Moritz et al., 2021). Even under the pressure of impending development, investigators still typically employ a Before-After-Control-Impact (BACI) or Beyond-BACI study design (Underwood, 1991, 1994) to establish changes occurring over time, even if observations are limited to only a few seasons. However, human stressors such as wastewater, dredging, and harvesting often predate the onset of even dedicated monitoring efforts (Todd et al., 2019). Demonstrating that shelly death assemblages are congruent with whole fauna, especially their pattern of compositional variation across a region as tested here, encourages their use as tools for estimating pre-monitoring (and potentially pre-impact) biological patterns. It should, of course, be remembered that sampling for death assemblages inflicts the same cost on living communities as if they were the target. The advantages of including death assemblages in monitoring efforts are becoming increasingly clear, however, namely in increasing the sample size of an important part of the benthic community, providing historical perspective on conditions during past decades and centuries, and as undertaken here, establishing the conditions (species:genus ratios, clade representation, evenness, etc.) under

which biologists in frontier areas might streamline monitoring to shelly clades alone.

Government-mandated, contracted, and non-profit monitoring programs might thus elect to integrate the collection and processing of death assemblages as part of their routine analytical procedures or, as exemplified here by cooperation between The University of Chicago and ECY, form an academic-agency partnership to explore an environmental issue of particular concern. ECY now saves rather than discards their sediment residue, allowing scientists at Chicago to develop counterpart ecological information from the death assemblage to tackle questions of mutual interest, such as this report. It also permits other applied and academic analyses, now underway, such as establishing scales of time averaging and uncovering environmental and species-level patterns of shell damage and loss, especially those relevant to ocean acidification. The University currently archives the growing library of dead samples from Puget Sound, as it does for counterpart material from cooperative agreements with municipal agencies in southern California (e.g., Bizjack et al., 2017; Tomašových and Kidwell, 2017; Leonard-Pingel et al., 2019). However, although agencies are typically required to retain live-collected material for a set number of years, they cannot accommodate the much larger volume of dead sieve residues, nor necessarily acquire funds for the labor of dead-shell analysis, at least until a formal protocol is established with clear benefits. Relationships with museums and other academic institutions are thus an important step. We strongly encourage academics to be proactive in proposing such collaborative efforts: agency personnel already have the necessary sampling operations in place and high-quality datasets on biological communities and environmental parameters, sometimes over impressive spatial and/or temporal scales as here and in southern California. Like all scientists, they like to see new value from existing data, especially from samples they had been discarding, and educational benefits are always important for public institutions. Additional mutual benefits – already experienced from this ECY-University of Chicago partnership – include (1) projects and connections for undergraduate and graduate students, who are also introduced to careers in applied science,

(2) publicity of conservation paleobiology and novel academic-agency partnerships to the public and stakeholders served by agencies, and (3) participation and authorship by agency scientists on publications and deliverables that appeal to a readership of both academic and applied scientists. Eventually, the ECY hopes that this cooperative effort, which extends beyond this first report on surrogacy, will reveal new insights that better inform their monitoring efforts, including the detection of human stressors that predate the PSEMP program, and encourage the implementation of new field methods (e.g., core sampling).

Although our focus here is on supporting modern-day efforts in conservation and environmental management, these results are also encouraging for paleoecological inference in older, pre-Holocene fossil records. The shelly component of a fauna cannot be used to directly reconstruct the composition of the unpreserved soft-bodied fauna, but tests like this support recognizing compositional variation of benthic communities more broadly, especially in the Cenozoic and late Mesozoic (i.e., the realm of the mollusk-rich Modern Evolutionary Fauna; Sepkoski, 1981; Rojas et al., 2021). By extension of death assemblages serving as a congruent proxy for the Modern whole fauna, fossil assemblages might reliably reflect patterns of compositional variation of the unpreserved fauna that was surely present. Surrogates of many types already have a long history of use in paleontology, for example using genus- rather than species-level data and focusing community analysis on a single class (as is also often done in community ecology, e.g., the dynamics of small mammals). Quantifying the confidence of such surrogates for application in the fossil record is thus an important direction for future research to better understand paleoecological variation among assemblages at basinal and larger regional scales.

1.6 Conclusions

This cooperative study between academic paleontologists and applied agency ecologists demonstrates that bivalve living and death assemblages are both reliable for recognizing

regional-scale compositional variation of the whole macrobenthic fauna of Puget Sound, a cold-temperate estuary with strong natural environmental gradients and potentially challenging taphonomic conditions. The success of surrogacy methods here, especially the suitability of family-level taxonomy and of using proportionally abundant (not necessarily rich) subsets of taxa, should promote the expansion of monitoring to new areas, where the costs and requisite taxonomic expertise needed for whole-fauna analysis might otherwise be discouraging. Such expansion can be an issue even in long-standing monitoring programs. We emphasize the importance of testing for the sensitivity of different surrogates locally prior to or at the onset of new monitoring programs, given variability in our results on numerical sufficiency (simplification of taxonomic count data) and in local motivations (e.g., surrogates to mirror the whole fauna versus identifying environmental gradients, or comparison with paleoecological evidence from death assemblages and cores). The power of bivalve death assemblages to detect regional compositional variation evident in the living whole fauna, despite potential taphonomic filters on composition and taxonomic identification, are especially encouraging, given the advantages of integrating paleoecological information with modern monitoring programs, a major thread in conservation paleobiology. Given that shelly death assemblages are already encountered by the sampling gear used for living macrobenthic infauna, they are a particularly inexpensive complement to long-term time-series data as well as a means of getting a baseline estimate where long-term data are lacking.

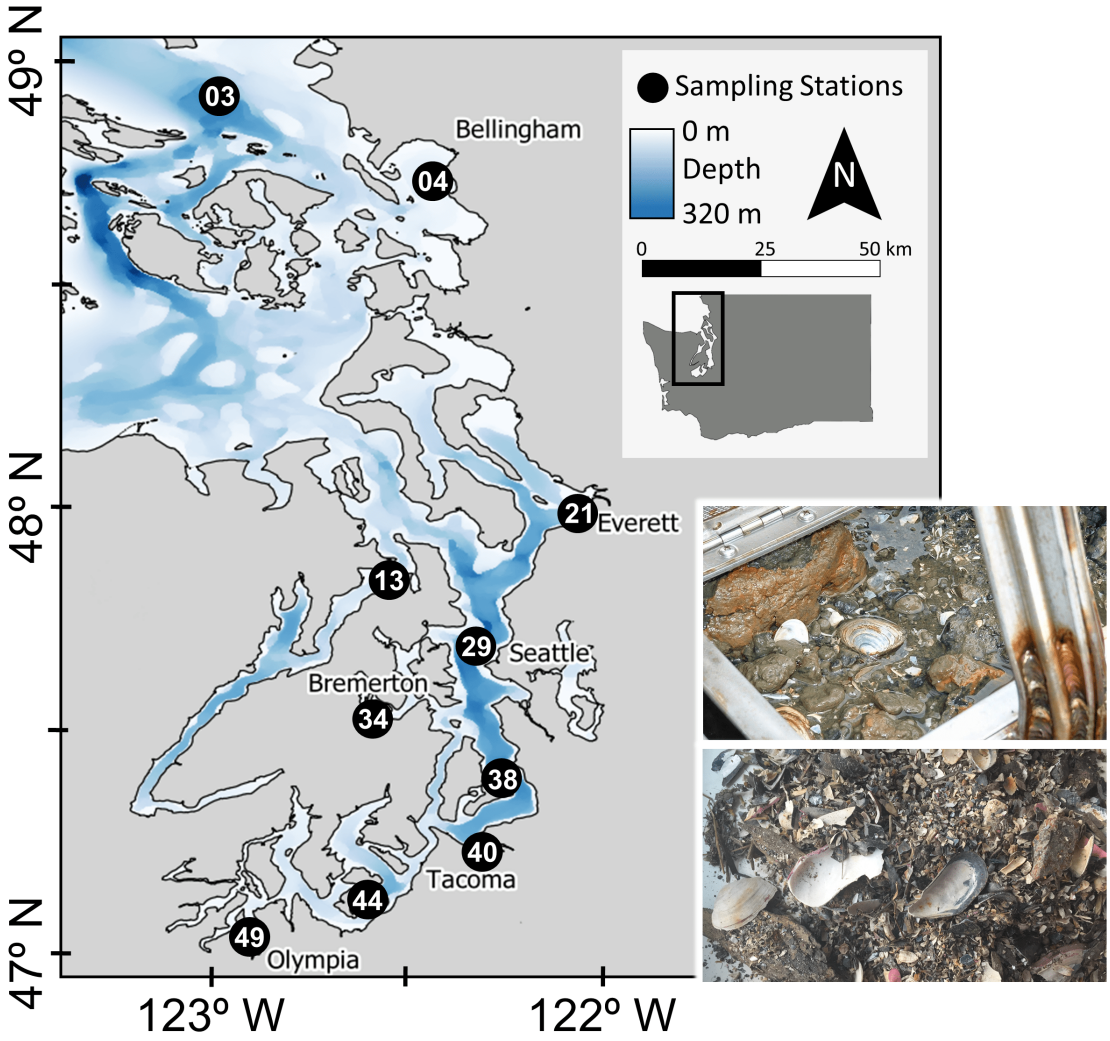

Figure 1.1: Map of the ten sampling station locations in the greater Puget Sound area, Washington State. Station identification numbers generally increase southwards. Inset images: a van Veen grab with sampled sediment and living organisms (top) and processed residue containing dead bivalve shells (bottom). Photos by ECY and B.S. Kokesh, respectively. Scale bar for dead shell photo = 5 cm.

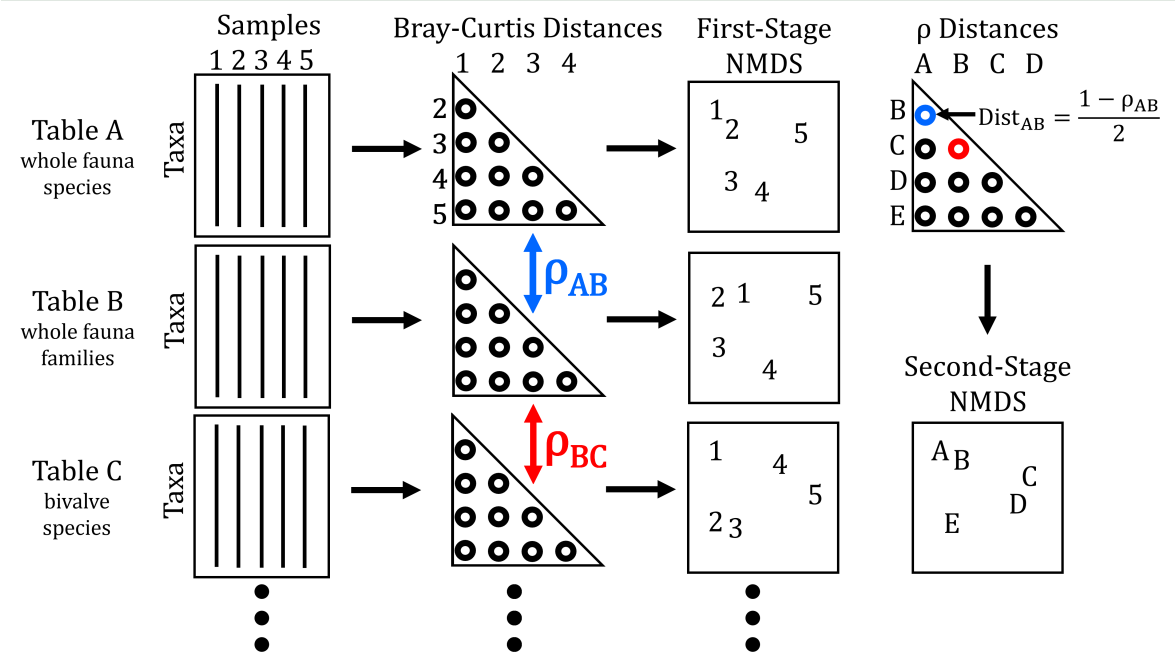

Figure 1.2: Schematic diagram of the procedure used to quantify inter-matrix correlations and produce second-stage NMDS among surrogate matrices. Starting with data tables for the same set of samples based on different combinations of taxa and numerical transformation, Bray–Curtis distance matrices are generated. These matrices can then be used to directly compute first-order ordinations (where points represent each sample, site, etc.) as typically performed. Correlation coefficients (ρ) calculated between each pair of distance matrices are computed into a distance between matrices (e.g., ρ_{AB} in blue is converted to $Dist_{AB}$). These distances are then assembled into a matrix used to compute second-stage ordination (where points represent different surrogate matrices). In the hypothetical example above, the locations of samples 1–5 in the first-stage NMDS plots for data tables A and B are more similar to each other than in C. Thus, the location of matrix C is strongly separated from matrices A and B in the second-stage NMDS plot. After Somerfield and Clarke (1995).

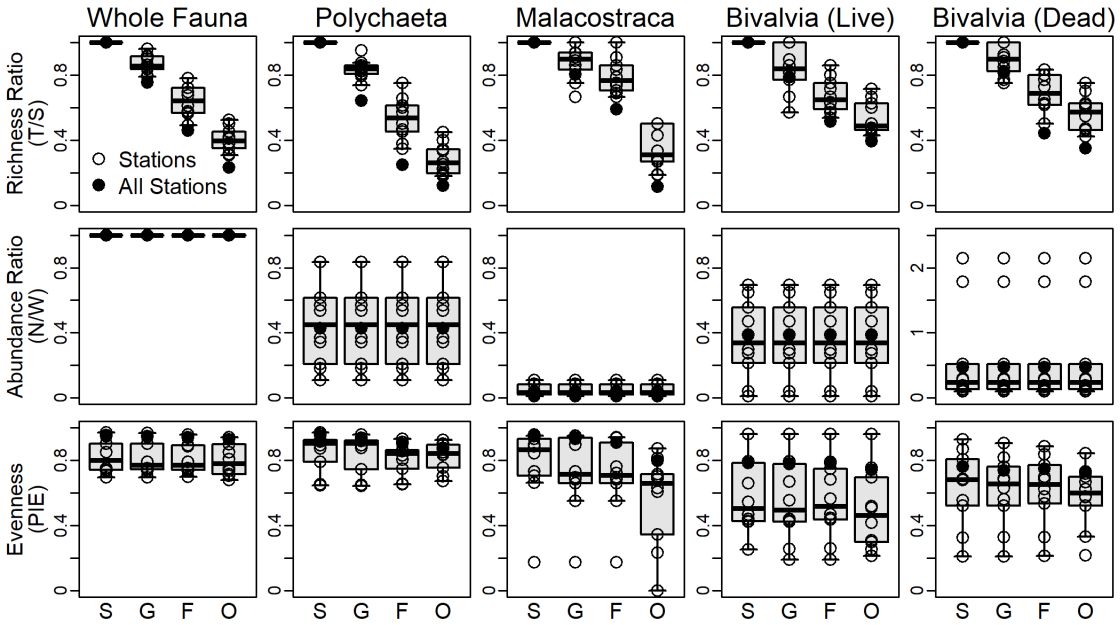

Figure 1.3: Univariate community metrics measured for surrogate assemblages at each station. Top row: richness ratios as the number of supraspecific taxa to the number of species (T/S). Middle row: abundance ratios as the number of individuals for each set over the number of individuals from the living whole fauna (N/W). Note the different vertical scale for dead bivalves. Bottom row: evenness as the probability of interspecific encounter (PIE) using proportional abundance data. Black circles are values calculated when data from all stations are aggregated. Boxplot whiskers extend to 5^{th} and 95^{th} percentiles. S, species; G, genera; F, families; O, orders.

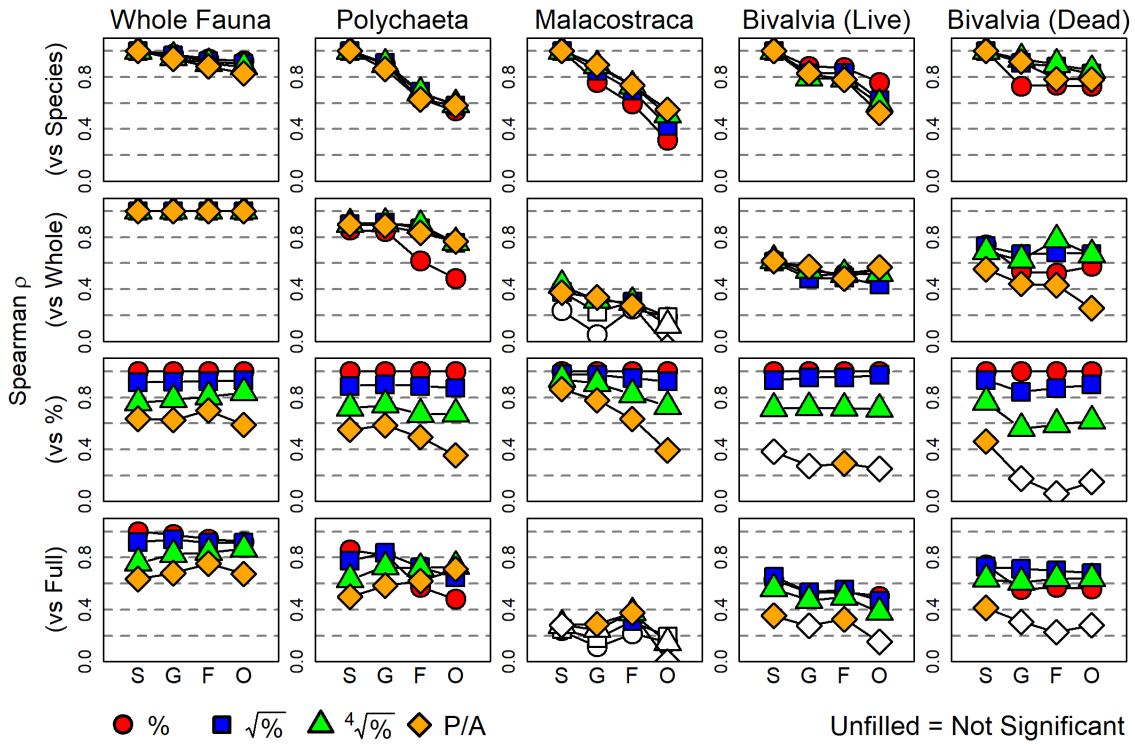

Figure 1.4: Spearman coefficients (ρ) for inter-matrix correlations representing the effect of each surrogacy method defined in this study. White-filled icons indicate correlations that were not significantly different from zero ($p > 0.05$). Top row: each matrix is compared to its species-level equivalent to test for taxonomic sufficiency. Correlations decline as the surrogate's resolution decreases. Second row: each matrix is compared to its whole-fauna equivalent to test for taxonomic subsetting. Correlations are strongest for polychaetes, moderate for bivalves, and poor for malacostracans. Third row: each matrix is compared to its proportional abundance (minimally-transformed) equivalent to test for numerical sufficiency. Correlation strength consistently decreases with increased severity of data transformation. Bottom row: each matrix is compared to the species-level whole-fauna dataset using proportional abundances (the "full" original dataset) to test for the combined effect of all three surrogacy methods. S, species; G, genera; F, families; O, orders.

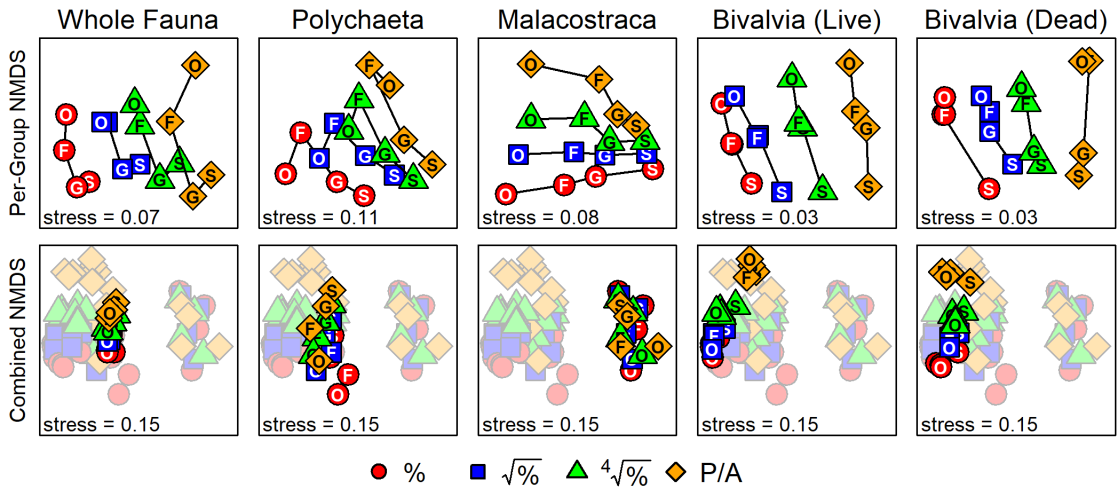

Figure 1.5: "Second-stage" NMDS ordinations of inter-matrix correlations based on Bray–Curtis distances. Top row: ordinations for each taxonomic set, demonstrating how matrices based on different taxonomic resolutions (species to orders) and data transformations (proportional abundances, %, to presence-absence, P/A) plot along two ordination axes. Bottom row: combined second-stage NMDS with surrogate matrices from all taxonomic sets. Matrices using the whole fauna cluster near those using polychaetes, with living and dead bivalves adjacent. Matrices using malacostracans strongly separate from the others. Data transformation severity also follows the second NMDS axis for all sets except malacostracans. The effect of taxonomic resolution on defining ordination space is weak, as matrices from the same set and based on the same transformation cluster tightly. S, species; G, genera; F, families; O, orders.

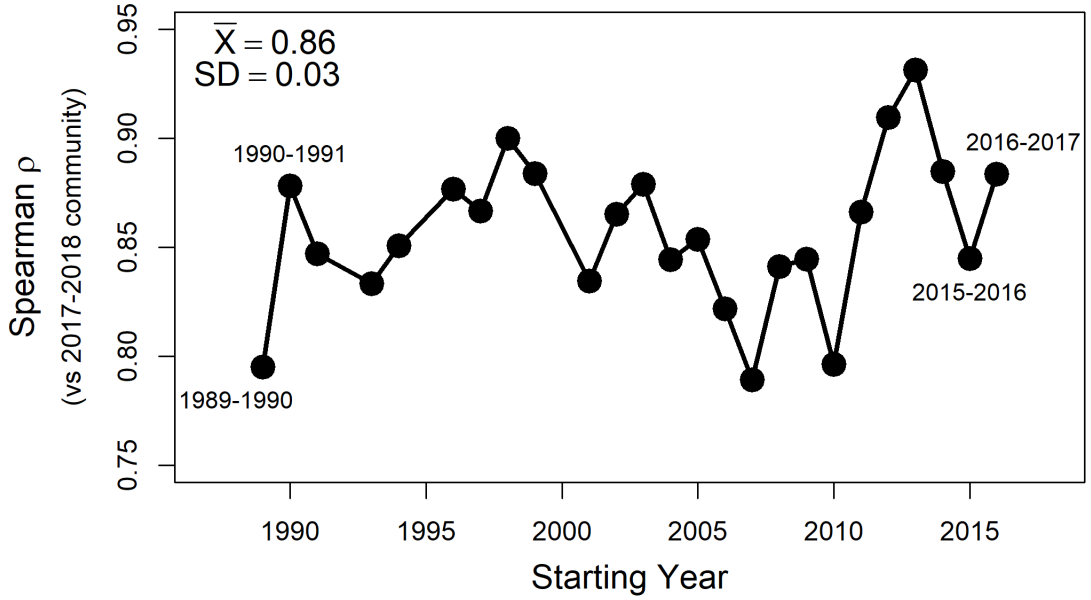

Figure 1.6: Spearman coefficients (ρ) for correlations between second-stage distance matrices using the 2017–2018 set of living assemblages and those using available 2-year sets from earlier in the 30-year community time series. Mean and standard deviation for the entire set of 2-year comparisons are in the upper left corner. Spearman ρ values are high ($\sim$0.8) across the entire time series, indicating that patterns generated the 2017–2018 dataset described in greater detail above have been consistent for the Puget Sound benthos over at least the past 30 years.

Table 1.1: Terminology of different types of taxonomic surrogacy as defined in this study.

Term	Definition
Surrogacy	Reducing the effort required to assess a biota
Sufficiency	Coarsening the detail/resolution/acuity of information
Numerical Sufficiency	Coarsening the measure of abundance (e.g., binary presence-absence or rank-abundances)
Taxonomic Sufficiency	Coarsening the taxonomic units (e.g., binning species into higher ranks or functional groups)
Subsetting	Narrowing the scope or breadth of information
Numerical Subsetting	Narrowing the number of specimens assessed (e.g., limit to first 100 individuals processed)
Taxonomic Subsetting	narrowing the taxonomic clades assessed (e.g., limit to a single class/phylum/functional group)

Table 1.2: Description of monitoring stations.

ID	Station Name	Water Depth (m)	% Fines	Latitude	Longitude
3	Strait of Georgia	227	76	48.87025	-122.9784
4	Bellingham Bay	26	96.2	48.68397	-122.5382
13	North Hood Canal	28	15.3	47.83758	-122.629
21	Port Gardner	25	56.6	47.98547	-122.2428
29	Shilshole	200	82.6	47.70075	-122.454
34	Sinclair Inlet	13	85.2	47.54708	-122.6621
38	Point Pully	201	95.5	47.42833	-122.3936
40	Thea Foss Waterway	10	33.1	47.2613	-122.4373
44	East Anderson Island	21	11.9	47.16133	-122.6736
49	Inner Budd Inlet	7	76.8	47.07997	-122.9135

Table 1.3: Summary of regional-scale macrobenthic community measures divided by taxonomic sets and resolutions. Metrics are the number of taxa (T), ratio of the number of taxa over the number of species (T/S), abundance of taxa (N), ratio of the abundance of taxa over the abundance of the whole fauna (N/W), and evenness, expressed as the probability of interspecific encounter (PIE). N and N/W do not change based on taxonomic resolution, so they are printed only once per set under the Species column.

		Species	Genera	Families	Orders
Whole Fauna	T	319	240	146	74
	T/S	1	0.75	0.46	0.23
	N	11,283			
	N/W	1			
	PIE	0.95	0.95	0.94	0.93
Polychaeta	T	157	101	39	19
	T/S	1	0.64	0.25	0.12
	N	4,817			
	N/W	0.43			
	PIE	0.95	0.94	0.9	0.87
Malacostraca	T	61	49	36	7
	T/S	1	0.8	0.59	0.11
	N	433			
	N/W	0.04			
	PIE	0.96	0.93	0.91	0.8
Bivalvia (Living)	T	33	26	17	13
	T/S	1	0.79	0.52	0.39
	N	4,367			
	N/W	0.39			
	PIE	0.78	0.78	0.79	0.74
Bivalvia (Dead)	T	43	35	19	15
	T/S	1	0.81	0.44	0.35
	N	5,337			
	N/W	0.47			
	PIE	0.76	0.74	0.75	0.73

Table 1.4: Summary of analysis of similarity results testing for the relative effect of different surrogacy methods on the variation of inter-matrix correlations among surrogates. Significant p-values are bolded.

ANOSIM test	R	*p-value*
Resolution	0.02	0.131
Subset	0.74	$<$**0.001**
Transformation	0.08	**0.003**

CHAPTER 2

DETECTING STRONG SPATIAL AND TEMPORAL VARIATION IN MACROBENTHIC COMPOSITION ON AN URBAN SHELF USING TAXONOMIC SURROGATES

2.1 Abstract

Surrogates of macrobenthic assemblages, intended to alleviate the effort and taxonomic expertise required for monitoring, can take many forms, such as using coarser taxonomic levels ('sufficiency') or only a subset of the whole fauna ('subsetting'). Here, the power of both approaches to retain community-level patterns of spatial and temporal variation were evaluated using an exceptionally long (47-year) infaunal dataset generated from monitoring wastewater impacts on an urban shelf in southern California. Four taxonomic sets (whole infauna, polychaetes, bivalves, malacostracans) were evaluated at five resolutions (species, genus, family, order, functional guild) along a pollution gradient subdivided into two spatial bins based on proximity to the wastewater outfall (near-field versus far-field) and three temporal bins based on wastewater treatment phases. All taxonomic sets detected weakening of the spatial gradient with improved wastewater treatment – communities became more similar in richness, evenness, and composition through time – and patterns were robust when coarsened to families or guilds. Polychaetes mirrored ('proxied') whole-fauna patterns most accurately, as expected since they constitute most of individuals and species. However, bivalves outperformed all other sets in detecting ('indicating') the pollution gradient itself, owing to their breath of feeding strategies. These results strengthen the consistently positive results from taxonomic coarsening emerging from tests elsewhere and the caveats for taxonomic subsetting: clade strengths serve different objectives. Comparable datasets should exist in environmental agency archives elsewhere, promoting the general surrogacy model. For monitoring programs still in their planning stage, regional insights could be acquired via

analogous nested analyses of a single survey.

This chapter was originally published as: Kokesh B.S., Kidwell S.M., Tomašových A., and Walther S.M. Detecting strong spatial and temporal variation in macrobenthic composition on an urban shelf using taxonomic surrogates. *Marine Ecology Progress Series*, 682:13–30, 2022b. doi:10.3354/meps13932.

2.2 Introduction

Intense urbanization of coastlines exerts an array of anthropogenic stresses on natural systems through resource harvest, landscape modification such as dredging and armoring, and the input of excess nutrients and other contaminants (Todd et al., 2019). These inputs include agrochemicals and fertilizers, industrial metals, and polycyclic aromatic hydrocarbons (PAHs) via stormwater, aquaculture, vessels, and municipal wastewater, with the latter usually constituting the largest volume of inputs (Islam and Tanaka, 2004). The environmental strain imposed by these human stressors is often evaluated by testing for before-after changes and/or far-versus-near (control-impact) contrasts in the composition and structure of benthic communities, which are sampled as part of sediment quality monitoring (e.g., Warwick, 1993; Gray and Elliott, 2009; Schiff et al., 2016). Such monitoring effort maintain a standardized quantitative sampling effort of both organismal abundance and diversity, covers a broad range of taxonomic clades, and identifies organisms to the species level or other fine-scale operational taxonomic units. Long-term programs extending a decade or more are unusual but provide invaluable insights into biological responses over the history of environmental stress (or stress alleviation) in a region (e.g., Borja et al., 2006; Stein and Cadien, 2009; Gogia et al., 2014; Clare et al., 2015; Borja et al., 2016; Schwing et al., 2017; Caswell et al., 2018). Such direct information is a kind of 'gold standard' for assessing the trajectory, and thus by implication the present status, of a community or ecosystem.

However, biological monitoring places a high demand on taxonomic expertise and labor,

especially in biologically diverse marine systems (Olsgard and Somerfield, 2000; Włodarska-Kowalczuk and Kędra, 2007). These expectations may potentially inhibit the establishment of new programs, especially in poorly studied regions where assessment of human impacts for public health and biodiversity conservation is urgent. In principle, these demands could be alleviated by using "surrogates" of species-level data on the whole fauna if the reliability of such surrogates in detecting the human impacts can be assumed (generalizing from studies elsewhere) and further validated. Beyond the concerns evaluating biodiversity and ecosystem health, the potential biasing effects of using surrogates is also important in historical ecology, which typically has to focus on the subset of taxa that are most likely to have been reported upon (e.g., commercial and sport taxa; Brind'Amour et al., 2014; Di Minin and Moilanen, 2014), and in paleoecology (our motivation), which typically has to rely on the subset of taxa with greatest postmortem durability under ordinary conditions of slow burial that characterize most marine environments (e.g., shelled mollusks, calcifying arthropods, and bryozoans; Schopf, 1978; Tyler and Kowalewski, 2017).

Surrogates fall into two categories: (1) using genera or other taxonomic levels coarser than species ('taxonomic sufficiency' *sensu* Ellis, 1985), and (2) using only a subset of the whole fauna, for example focusing on a single class ('taxonomic subsetting'; Table 2.1). Both reduce the labor and time required to process a sample after sampling, but are rarely evaluated together.

Taxonomic sufficiency has been the subject of innumerable tests in marine settings, mostly finding that coarser taxonomic resolution suffices to detect both anthropogenic (Warwick, 1988; Ferraro and Cole, 1992; Buss and Vitorino, 2010; Dimitriou et al., 2012; Stark et al., 2014) and natural (Włodarska-Kowalczuk and Kędra, 2007; Bevilacqua et al., 2009) spatial gradients (for freshwater examples, see Jones, 2008 and Mueller et al., 2013, and for terrestrial invertebrates, see Pik et al., 1999, and Souza et al., 2016). Non-Linnaean classifications such as morphospecies and functional guilds have also demonstrated promising

results (Bhusal et al., 2014; Brind'Amour et al., 2014).

Tests of taxonomic subsetting have had more diverse objectives (Moreno et al., 2007). In most studies, authors sought to identify a single clade either able to mirror variation in richness or some other attribute of the whole fauna (i.e., a "proxy" for the whole fauna; e.g., Olsgard et al., 2003; Magierowski and Johnson, 2006; Fontaine et al., 2015) or able to efficiently detect a pollution or other gradient (an "indicator" group), even if it exhibited poor congruence with the whole fauna (Magierowski and Johnson, 2006; Heino, 2010). Subset surrogates thus fall into one of two types: 'proxy subsets' (good for mirroring the whole fauna) and 'indicator subsets' (superior at detecting a particular environment or perturbation; Table 2.1).

How effective are both taxonomic sufficiency and subsetting modes of surrogacy, as well as their interaction (i.e., subsets at multiple resolutions) at detecting spatial and temporal gradients in anthropogenic disturbances? Here, an exceptionally long-term (∼50 years) and taxonomically consistent species-level macrobenthic database was used to test whether surrogates upheld patterns observed using the whole fauna across a well-known anthropogenic pollution gradient over successive phases of wastewater abatement. The dataset, which consists of annual infaunal grab samples from 11 stations in 61 m water depth along the Palos Verdes shelf in southern California since 1970, encompasses about 600 species from 11 phyla. Previous analyses of the full infaunal dataset, including municipal annual reports, have demonstrated improvement in benthic conditions associated with enhanced wastewater treatment (Stein and Cadien, 2009; LACSD, 2020).

This analysis takes opportunistic advantage of a dataset generated under a government regulatory requirement to conduct long-term monitoring of benthic community health and recovery in response to contamination reductions, and thus was not designed as a formal research experiment. In contrast to a standard Before-After-Control-Impact "BACI" study design (Underwood, 1991; 1994, and references therein), here the Before-After component is

a temporal analysis over a history of wastewater effluent pollution reduction (Phase 1 with only primary treatment, Phase 2 with partial secondary treatment, and Phase 3 with full secondary treatment). Wastewater discharges to the Palos Verdes shelf began in the 1930s and regulation only in the early 1970s, and thus benthic response to the onset and rise of pollution is not known directly (but see a sediment-core based analysis of bivalve shells; Leonard-Pingel et al., 2019). The Control-Impact component here is a spatial comparison of far-field ("reference") and near-field areas on the shelf relative to the pipe opening, with those two areas defined by agency scientists on the basis of sediment chemistry in the 1970s. Formal controls do not exist: no stations on the shelf have been free of pollution for the entire duration of monitoring. The disadvantages of this necessarily less stringent design are, we believe, outweighed by some unusual advantages, namely (a) an exceptionally long time series (47 years) with (b) scrupulous taxonomic consistency despite turnover in agency personnel (SCAMIT, 2021), (c) a strong expert-knowledge base for species biology in the region (Coan et al., 2000; Macdonald et al., 2010; SCAMIT, 2021), (d) consistent, multiple sampled sites in both near- and far-field areas over the entire duration, (e) consistent sampling methods, and (f) holding water depth constant.

2.3 Methods

2.3.1 Study area

The Palos Verdes shelf in southern California has a long history of receiving wastewater contaminants beginning with the 1937 opening of an ocean outfall in 34 m-deep water at White Point; two Y-shaped outfalls dispersing wastewater at 60 m depth were later installed in 1956 and 1966 (Figure 2.1). The 61-m outfalls remain in use and wastewater treatment processes have continually improved since 1970 in anticipation of the passage of the US Clean Water Act in 1972. Wastewater was initially subject only to primary treatment, and emission

rates steadily increased with 20^{th}-century urbanization, with total suspended solids loadings peaking at ∼150,000 t/yr (Figure 2.2). Emissions of solids (to which most contaminants adhere) began to decrease markedly after the onset of advanced primary treatment in the 1970s, with additional declines in solids and in biological oxygen demand following the onset of partial secondary treatment in the 1980s and 1990s and full secondary treatment starting in the 2000s. Despite a high volume of water outflow to the ocean over the last 50 years, contaminants in the effluent are now well below allowable emission limits (LACSD, 2020).

This emission history created a strong point-sourced spatial gradient in sediment contamination (Swartz et al., 1991; Eganhouse and Pontolillo, 2000) that extends mostly northwest of the outfall area, deflected by the Southern California Countercurrent (Figure 2.1). The sediments associated with the effluent discharge form a fine-grained, organic-rich mound as thick as 60 cm near the outfall and include numerous contaminants including DDT and PCBs (Stull et al., 1996; Lee et al., 2002). This contaminated sediment layer thins with distance from the outfall; stations where a contaminated sediment layer was detected are designated as near-field stations and those without such a layer are denoted as far-field (not formal controls, because they likely experienced direct or indirect effects from contaminants such as excess nitrogen in the fine sediments released prior to the onset of refined treatment technologies; LACSD, 2020). Along the 61-m isobath today, sediment sampled from the immediate near-field area is chemically undistinguishable from that in the far-field stations to the northwest and southeast (Stein and Cadien, 2009; LACSD, 2020).

Long-term monitoring of macrobenthic communities using benthic grabs were conducted on the Palos Verdes shelf by the Los Angeles County Sanitation Districts (LACSD) starting in 1970 and are ongoing. Their sampling grid comprises 44 stations: 11 bathymetric transects each sampled at 30, 61, 152, and 305 m water depths (Figure 2.1). Infaunal and other data from this standardized effort have revealed a significant and positive benthic response to improved wastewater treatment, including increasing macrofaunal species richness and

evenness (LACSD, 2020). The faunal composition has, throughout the grid, reached or increasingly approached a reference condition as defined by the Benthic Response Index (BRI), a metric based on the pollution tolerance and abundances of local species calibrated to regional disturbance gradients (Smith et al., 2001), and as compared to other areas within the Southern California Bight that are considered to be undisturbed.

2.3.2 Data preparation

LACSD sampled macrobenthic invertebrate communities from their full spatial grid from 1972 to the present, with the dataset used here encompassing 1972-2019. Samples were collected semi-annually in February and August until 2006, after which the agency switched to annual sampling during the summer. Samples were initially collected using Shipek grabs, which were replaced by van Veen grabs after 1980; here, Shipek grab data have been transformed to equal the volume of van Veen grabs by pooling three replicate Shipek grabs taken at each station. Sediments were sieved through 1-mm mesh with seawater on deck, fixed in 10% formalin, and then preserved in ethanol.

Agency biologists sorted and identified all invertebrates to the lowest level possible, e.g., species. Taxonomic identities have consistently been standardized to a regional listing compiled and regularly updated by the Southern California Association of Marine Invertebrate Taxonomists (SCAMIT, 2021), a volunteer group of regional professionals; the SCAMIT listing is used by all regional municipal agencies and in the Southern California Bight Regional Monitoring Program (e.g., Schiff et al., 2016). All species assignments in the dataset have been updated to a single, internally consistent taxonomy (SCAMIT, 2021) for the present analysis.

The analyses for this study focused only on samples from the 61-m isobath, which coincides with the effluent ports of the outfall that have been active point-sources since the 1950s. Each of the 11 stations along this middle-shelf isobath was categorized as either "near-

field" (stations 5-8) or "far-field" (stations 0-4, 9-10) based on sediment chemistry conditions during the first decade of monitoring (Stull et al., 1996), an assignment that reflects physical proximity to the outfall and the prevailing northwest-directed California Countercurrent (Figure 2.1). Samples collected in February during years 1972-2005 were excluded, yielding a dataset that consisted of one summer-time sample per station per year from 1972 through 2019 for a total of 1,939 samples.

This 47-year history was subdivided into three temporal phases based on the treatment level of wastewater emissions: primary and advanced primary (Phase 1 in 1972-1983), partial secondary (Phase 2 in 1984-2002), and full secondary (Phase 3 in 2003-2019; Figure 2.2). At each station, samples were pooled by these intervals to create three temporal bins (referred to as 'phases' below). The final dataset thus consisted of three temporal bins at each of 11 stations, which were assigned to one of the two outfall categories (spatial bins, referred to as near-field and far-field 'areas' in Results). Although seabed conditions clearly improved throughout the Palos Verdes sampling area over these decades of improving wastewater treatment (LACSD, 2020), a decision was made to leave stations in their outfall categories as originally defined based on conditions in 1972, allowing us to test the ability of surrogates to detect weakening of the spatial gradient.

Given the stepwise decrease of emissions across the boundaries of treatment phases (Figure 2.2), the robustness of datasets to possible lags in benthic response were tested by omitting the initial one or two years of data immediately following a change in treatment. Omitting these data did not affect results significantly, and so all years of data were retained in analyses.

2.3.3 Data analysis

All data extraction and analyses were conducted using R version 4.0.3 (R Core Team, 2023). The original dataset was transformed to 20 surrogate matrices (henceforth called 'sets') for

coarsened taxonomic levels and selected taxonomic subsets. First, species-level counts for the whole fauna were aggregated into genus, family, order, and functional guild level counts (tests at five levels of taxonomic resolution). Species were assigned to functional guilds (simply 'guilds' hereafter) following Macdonald et al. (2010), who grouped East Pacific species on the basis of animal motility, life habit, food source, diet, and feeding mode. Species in our dataset that were not listed by Macdonald et al. (2010) were omitted from guild analyses. Although guild classifications are often found to be highly conserved within and among related clades (e.g., bivalve families; Stanley, 1970; Collins et al., 2019), they should not be viewed as equivalent to Linnean ranks. Using the assignment data from Macdonald et al. (2010), 33 (10%) of the total 335 families in common with Macdonald et al.'s list contained more than one guild (Table 2.2). However, no single genus in our dataset contained more than one guild. Thus, our analysis of guilds represents a coarsening of resolution between the levels of genus and family, but that coarsening is only slightly less than that represented by families.

After taxonomic coarsening the species-level data into these five ranked-based matrices, new matrices were extracted from each for the three most abundant classes represented, namely Polychaeta, Bivalvia, and Malacostraca (three taxonomic subsets, to be compared with the whole-fauna set). These three clades together represent ∼90% of individuals in the whole-fauna (Figure 2.3). We chose to assess this many subsets in order to determine whether some behaved as proxy subsets and others as indicator subsets (Table 2.1). All procedures that follow were applied to each of these 20 matrices – five resolutions of information for three taxonomic subsets and for the whole fauna.

The distribution of the number of species within the five supra-specific categories (i.e., groupings higher than species level) was quantified as the ratio between the raw number of species (i.e., richness S) over the number of genera (G), families (F), orders (O), and functional guilds (FG) (e.g., S/G, S/F, S/O, S/FG). The smallest possible value of any ratio

is 1, which indicates that the higher taxon is monospecific.

Richness and evenness were compared between near-field and far-field samples for each treatment phase by comparing the median values and interquartile range (IQR) of one area to the other. Richness was rarefied to the smallest sample size per taxon per treatment phase to account for unequal abundances among the taxonomic sets (3,727 individuals for the whole fauna, 1,783 Polychaeta, 285 Bivalvia, 44 Malacostraca). Evenness is expressed as the probability of interspecific encounter (PIE; Hurlbert, 1971), which has low sensitivity to sample size and richness. PIE values have potential to range from 0 to 1, with 0 indicating a sample where all individuals are from a single taxon (i.e., the chance of two randomly-selected individuals being different taxa is zero), and with a PIE value of 1 indicating that individuals are equally distributed among all taxa.

Compositional variation among samples was assessed using non-metric multidimensional scaling (NMDS) based on Bray-Curtis distances computed with square-root transformed proportional abundances (Hellinger transformation). Bray-Curtis values quantify the compositional dissimilarity among stations on a scale from 0 to 1 (Bray and Curtis 1957). NMDS ranks these distances prior to ordination on a fixed number of axes (here, k = 2). Compositional differences among near-field and far-field stations for each treatment phase were also tested using permutational multivariate analysis of variance (PERMANOVA; Anderson, 2001). R^2 values, generated by PERMANOVA, which measure the amount of variation explained by sample categorization of near-field and far-field conditions, were compared among sets to assess the relative strengths of surrogates to delineate the near-field region of the shelf during each phase of wastewater treatment.

The contributions of individual taxa to the Bray-Curtis dissimilarity between near-field and far-field communities for each set were calculated using the similarity percentage procedure (SIMPER; Clarke, 1993). To quantify the proportion of taxa that contribute most to dissimilarity for each set, the number of taxa that met increasing levels of cumulative

contribution (25%, 50%, and 75%) was divided by the total number of unique taxa. Higher proportions at contribution thresholds indicate that a higher number of taxa are needed to account for the observed dissimilarity.

2.4 Results

2.4.1 Taxonomic structure of the macrobenthic fauna

The 'whole fauna' dataset comprised 536,056 individual organisms across 1,277 species, 722 genera, 347 families, 120 orders, 34 classes, and 13 phyla, and these species encompass 86 guilds (Table 2.3). The three focal classes comprised 89% of all individuals, with polychaetes encompassing 61%, bivalves 17%, and malacostracans 10% (left panel; Figure 2.3). These three classes also comprised 72% of all species: 42% for polychaetes, 7% for bivalves, and 23% for malacostracans (right panel; Figure 2.3). Seven of the ten most abundant classes were also the highest in species richness.

Polychaetes were the most speciose per higher taxon and bivalves were the least speciose subset (Table 2.3). For example, the species/genus (S/G) ratio for polychaetes was 2.2 as opposed to 1.8 for the whole fauna, and 1.4 for bivalves. Per-family species richness for polychaetes (S/F = 11) was ∼3 times higher than other subsets and the whole fauna (S/F = 3.1-3.7), and per-order species richness of both polychaetes and malacostracans was ∼3 times higher than the whole fauna (S/O = 33.7, 36, and 10.6, respectively). Bivalves were the least speciose subset at all levels, with the contrast increasing at successively higher levels (Table 2.3).

Considering the whole fauna, guilds included larger numbers of species than did orders (S/FG = 14.9, S/O = 10.6; Table 2.3). In each taxonomic subset, the S/FG ratios typically fell between the S/F and S/O levels. The polychaete S/FG ratio (15.8) most closely resembled that of the whole fauna (14.9), while those of bivalves and malacostracans were notably

smaller (S/FG of 7.4 and 9, respectively).

2.4.2 Temporal and spatial variation in richness

Temporally, rarefied species richness increased with improved wastewater treatment in all four taxonomic sets – the whole fauna and three class-level subsets (first column in Figure 2.4). Richness of the whole fauna, polychaetes, and bivalves exhibited stepwise increases across each phase boundary; a gain of ∼100 species by the whole fauna, ∼50 by polychaetes, and ∼10 by bivalves. Malacostracans exhibited a large increase from Phase 1 (1972-1983) to Phase 2 (1984-2002), but negligible change from Phase 2 to Phase 3 (2003-2019).

Spatially, with the exception of the bivalve subset, species richness was lower at near-field than at far-field stations, with the contrast strongest during Phases 1 and 2 and smallest during Phase 3 (Figure 2.4). Bivalves exhibited no difference in richness between near-field and far-field stations during any of the treatment phases.

These species-level differences in richness over time and between areas persisted with taxonomic coarsening to the family level, although family-level polychaete data dampened the magnitude of temporal increase across each phase boundary. Coarsening data to the ordinal level preserved patterns for the whole fauna, but significantly alters patterns for each subset, particularly for malacostracans (fourth column in Figure 2.4). Guilds appear to have the opposite effect from orders: using guild-level data altered patterns for the whole fauna, but preserved patterns, albeit slightly dampened, for each of the subsets (fifth column in Figure 2.4).

2.4.3 Temporal and spatial variation in evenness

With improved wastewater treatment, species-level evenness generally increased and variation among stations within both the near-field and far-field areas generally decreased (i.e., smaller IQRs); these changes were observed in all sets except malacostracans, which main-

tained consistently high evenness (Figure 2.5). Bivalves exhibited the strongest increase in evenness over time – stronger than the whole fauna – and polychaetes most closely resembled the whole fauna, but all sets reached an evenness >0.9 by Phase 3.

Spatially, during each treatment phase, polychaetes mirrored patterns exhibited by the whole fauna with no significant difference in evenness between near-field and far-field stations. In contrast, bivalves had far higher evenness at near-field than far-field stations during both Phase 1 and Phase 2 (Figure 2.5). By Phase 3, these differences among bivalves disappeared, with evenness values from both near-field and far-field stations reaching 0.9-1.0.

As with richness, species-level patterns persisted with coarsening of taxonomic resolution up to the family level (left to right across Figure 2.5). Ordinal-level data did not change patterns in the whole fauna but did alter polychaete and bivalve patterns (mostly a decline in maximum evenness). Malacostracans had greater ranges of evenness and a particularly strong temporal decline at the ordinal level. Guilds preserved similar patterns as the family level for the whole fauna, polychaetes, and bivalves, but slightly altered malacostracan patterns.

2.4.4 Temporal variation in strength of spatial gradient

At the species level, ordination revealed a consistent pattern of temporal and spatial separation among samples (Figure 2.6; each icon is a station within either the near-field or far-field area during a treatment phase). Generally, successive water treatment phases separated along NMDS axis 1, while near-field and far-field stations within each phase separate along NMDS axis 2. All ordinations revealed a decrease in multivariate dispersion over time – smaller convex hulls; i.e., less compositional variation among sampled stations within an area – as well as progressive convergence in community composition of the near- and far-field areas (a decrease over time in beta diversity along the 61-m isobath).

These species-level patterns persisted for all sets up to the family level and also for guilds. Coarsening resolution to the ordinal-level changed the relative positions of samples,

but temporal translation and spatial homogenization are still apparent.

These temporal changes were also detected by PERMANOVA. Near- and far-field areas differed significantly in composition during Phase 1 and less strongly over successive treatment phases, becoming insignificant during Phase 3 (Figure 2.7). Together with NMDS (Figure 2.6), PERMANOVA results indicated spatial homogenization in community composition and a loss in strength of the spatial pollution gradient over time. R^2 values for polychaetes and malacostracans were similar to the whole fauna, progressively declining from $\sim$0.45 during Phase 1 to 0.3 during Phase 2 and $<$0.2 during Phase 3. Bivalves demonstrated both the strongest early differences ($R^2 = 0.65$ during Phase 1) and the steepest decline across treatment phases (0.4 during Phase 2, $<$0.1 during Phase 3). Resolution had little impact on PERMANOVA results up to the family level and for guilds: at the ordinal level, the whole fauna pattern was unaffected, polychaetes detected significant differences during Phase 3, bivalves detected even stronger differences during Phase 1 and Phase 2, and malacostracans no longer detected differences during Phase 2. Ordinal-level data thus significantly affected PERMANOVA results.

SIMPER identified that two species are most responsible for spatial variation in the composition of the whole fauna during Phase 1, namely the opportunistic polychaete *Capitella capitata* (9% contribution) and the chemosymbiont-bearing bivalve *Parvilucina tenuisculpta* (7% contribution). All other species had $<$4% contribution. Abundances of *C. capitata* were much higher in the near-field area during Phases 1 and 2 while *P. tenuisculpta* was more abundant in the far-field area (Figure 2.8). Both species declined exponentially in numerical abundance over successive wastewater treatment phases (from densities $>10^3$ to $<10^2$ for both species). By Phase 3, neither species was a dominant contributor to spatial dissimilarity, which had also become insignificant for the whole fauna (white bars in Figure 2.7).

Richness increased from Phase 1 to Phase 2 for all sets at all taxonomic resolutions but

remained mostly stable from Phase 2 to Phase 3 (Total S; Table 2.4). However, SIMPER for the whole fauna, polychaetes, and bivalves indicate that spatial homogenization of communities – i.e., the weakening of spatial gradients – continued from Phase 2 to Phase 3 (parenthetical values; Table 2.4), with decreasing per-taxon contributions to dissimilarity across all three treatment phases. These patterns persisted with coarsening taxonomic resolution up to the ordinal level. Patterns for malacostracans were less consistent: while the number of sub-ordinal taxa increased during each phase, the proportion of taxa that notably contributed to dissimilarity decreased.

2.5 Discussion

A vast number of analyses on taxonomic surrogacy have been conducted in the past, spanning numerous locations, environmental conditions, and faunas (Moreno et al., 2007; Bacci et al., 2009). This wealth of information and some meta-analyses (Mellin et al., 2011; Westgate et al., 2014) indicate that the usefulness of surrogates depends on the study system (e.g., substrate, latitude, strength of the environmental gradient, spatial scale), the study aim (e.g., detection of gradients, estimation of diversity, richness hotspots), and the methodologic focus (taxonomic sufficiency, which tends to yield positive results from coarsening resolution, versus taxonomic subsetting, which yields highly varied results).

The Palos Verdes shelf evaluated here is a single study system where surrogates might be unusually effective. However, our multi-decadal time-series of annual sampling with exceptional taxonomic consistency permits an exceptional breadth of tests without the challenges of comparing among published studies. Comparable datasets might well exist in the archives of environmental management agencies elsewhere, permitting counterpart analyses to support a general model of surrogacy (e.g., shelves with differing productivity, tropical shelves, coastal embayments). In addition to spurring the discovery and academic re-use of such datasets, nested analyses similar to those here could be applied to a pilot survey, allow-

ing managers to evaluate the viability of potential surrogates to satisfy a new monitoring program's objectives.

2.5.1 Caveats

The dataset evaluated is itself only a subset of the total macrobenthic fauna in the study area owing to the use of Shipek and van Veen grab samplers, which are most effective at collecting infaunal taxa living within the upper ∼15 cm of the seabed (minimum penetration required to retain a sample; LACSD, 2020). Although these data largely exclude mobile epifauna, and certainly planktonic meiofauna, the taxa collected by these and other point-penetration sampling gear (e.g., 0.01 m^2 box cores) are the basis of most quantitative analyses of soft-sedimentary communities by both academics and agencies (Eleftheriou and Moore, 2013). It is unknown how results may vary if we instead or additionally employed sampling methods (trawling, plankton nets, etc.) that better target these other commonly-studied aspects of the overall community.

Numerical sufficiency (Table 2.1), although not assessed here, can significantly reduce the effort of counting individuals in samples (e.g., Carneiro et al., 2010; Landeiro et al., 2012) but limits the analytical utility of data. Presence-absence, ranked abundance, and categorical (abundant, common, rare, etc.) data, for example, will quantify raw richness accurately, but cannot be rarefied and are unsuitable for evaluating evenness (community structure). Such low-resolution data are also problematic for SIMPER analysis, which specifies the use of abundance-based Bray-Curtis dissimilarities (Clarke, 1993). For these reasons, the absolute counts of individuals for each sample – i.e., density data – were used in all analyses.

Data transformation (i.e., analytical coarsening) of numerical abundance data constitutes a test of numerical sufficiency, but of course does not reduce processing effort and thus is not a surrogate. Up to the family level, our patterns using Hellinger-transformed data are robust to other transformations of proportional abundances, including no transformation or

fourth-root transformation. The main effect of data transformation is on the strength of the spatial difference detected by taxonomic subsets during Phase 1. Although differences are significant during Phase 1 for all surrogates, PERMANOVA R^2 values decline as the transformation intensity increases; the same trend was observed in the other temporal bins. Coarsening numerical resolution to the extreme of presence-absence data makes the patterns robust only to the genus level.

We use the three most numerically abundant macrofaunal classes as subsets in this study, but it should be noted that the class Gastropoda ranks slightly above Bivalvia in species richness in this dataset (Figure 2.3). However, gastropods have such low abundance on the Palos Verdes shelf that they did not, on their own, detect any spatial or temporal variation. Thus, in serving as surrogates, taxonomic subsets represented by large numbers of individuals out-perform subsets with large numbers of species. The analyses in this study thus focused on the three most numerically abundant, rather than most speciose, classes as taxonomic subsets.

2.5.2 Taxonomic resolution: genus- and family-level data suffice

Coarsening taxonomic resolution to genera and families preserved both the spatial and temporal gradients in wastewater contamination that were detected by the whole macrobenthic fauna using species-level data, and the same was true for all assessed surrogate subsets, whereas ordinal-level data only detected a spatial gradient when one existed (i.e., during Phases 1 and 2; Figure 2.7). These results corroborate previous studies that taxonomic coarsening to family-level identification is a good first-order means of reducing the effort of detecting spatial gradients (e.g., Jones, 2008; Heino, 2010; Mellin et al., 2011; Pitacco et al., 2019). Our analysis explicitly extends this result to temporal variation and, more importantly, reveals its (mostly stable) interaction with different taxonomic subsets of the fauna (see sections 2.5.3 and 2.5.4).

Genus- and family-level data probably sufficed in this southern California setting because S/G and S/F ratios in the whole fauna are relatively low (∼2 and 5, respectively; Table 2.3), consistent with studies in other systems (e.g., Bevilacqua et al., 2012; de Oliveira et al., 2020). Sufficiency is commonly found to be inversely related to the ratio of the number of species over the number of higher taxa (S/T): increasing the number of species per higher taxon increases the probability that these species will differ in their pollution tolerance. Some studies postulate that S/T values <2.5 will generally result in good correlations with species-level patterns (Timms et al., 2013; Albano et al., 2016). Although the family and guild ratios in this study are almost all >2.5 (Table 2.3), the S/T values on the Palos Verdes shelf corroborate the growing evidence that the sufficiency of higher taxa depends on S/T ratios remaining below a particular threshold.

The unpredictable patterns produced by ordinal data across all faunal datasets probably arise at least in part from the high S/O values (8.9-36, Table 2.3), although the small number of orders probably also contributes to their low sensitivity. In malacostracans, for example, where diversity and compositional patterns deteriorate most strongly with coarsening resolution, the S/T ratio rises strongly from families (S/F = 3.2) to orders (S/O = 36).

The exceptionally strong anthropogenic environmental gradient along the Palos Verdes shelf during Phases 1 and 2 also favors the practicality of coarser groupings: habitats as distinct as black H2S-rich mud and greenish sandy mud are all soft but are likely to be dominated by different guilds rather than simply by different species within a single, consistently dominant guild. Guilds have been shown to be evilutionarily-conserved within families and genera (for bivalves, see Stanley, 1970; Collins et al., 2019). Consequently, coarsening taxonomic resolution to guild-level information should generally not reduce the ability to differentiate spatial and temporal variation, especially in areas with known or suspected strong human stressors.

Finally, species-level taxonomy undergoes revision, even in areas with well-resolved macrobenthic systematics such as southern California, challenging the merging of species-level datasets across decades of monitoring. The recognition of new species by splitting older ones creates the greatest complications for historic compilations in the absence of voucher material retained for future reference. For example, in biological monitoring of California shelves, the bivalves *Tellina carpenteri* and '*Tellina sp B*' were not differentiated prior to the 1990s. Such newly split taxa will increase the S/G ratio, but they constitute a very small minority of taxa in the fauna. Other taxonomic revisions are, in contrast, easily applied retrospectively to data and will tend to decrease the S/G ratio, such as the correction of synonyms and the upgrading of subgenera to genera. The greater challenge is the per-genus species richness hidden by known but unresolved taxonomic complexes, that is, taxa that harbor cryptic species whose distinction is beyond conventional morphologic detection (e.g., the polychaetes *Capitella capitata* and *Scoloplos armiger*). Fortunately, such complexes constitute, again, only a small number of taxa in the temperate southern California fauna. A potential solution to alleviating such challenges, especially in comparisons among different regions, would be the implementation of environmental DNA (eDNA) metabarcoding to identify operational taxonomic units (OTUs) without the biases of phenotypic taxonomic identification (Ruppert et al., 2019). The ecological anonymity of OTUs will, however, represent a significant loss of insight.

2.5.3 Polychaetes are the best proxy of spatial and temporal variation in the whole fauna

Polychaete patterns resembled the whole fauna most closely, indicating that they are the best proxy subset (Table 2.1) for characterizing spatial and temporal variation of benthos on the Palos Verdes shelf. This fidelity likely arises from polychaetes constituting the majority of individuals in the whole fauna (61%; Figure 2.3; Table 2.3) – not simply a plurality –

and persisted with coarsening to the genus and family levels (Figures 2.4-2.7). In contrast, constituting a high proportional richness of the fauna did not make a clade here an inherently effective proxy (e.g., malacostracans and the species-rich but individual-poor gastropods). Thus, although polychaetes are the most speciose class in our fauna, their proxy power results from their high proportional numerical abundance.

For similar reasons, other studies have also found polychaetes to be excellent proxies of the whole fauna (e.g., Giangrande et al., 2005; Magierowski and Johnson, 2006; Włodarska-Kowalczuk and Kędra, 2007). Subsets finer than a class, however, are highly variable in power. For example, Olsgard et al. (2003) identified the order Terebellida as an effective proxy subset for all polychaetes and, to a lesser extent, the whole fauna in a spectrum of temperate to subarctic shelves. In our California 61-m isobath dataset, Terebellida was the most numerically abundant order (34% of polychaetes and 21% of the whole fauna), but failed to mirror spatial and temporal patterns of either the full polychaete subset or the whole fauna; this family also failed to detect the spatial pollution gradient. Capitellidae, the second most abundant polychaete clade (23% of polychaetes, 14% of whole fauna), also failed to mirror the full polychaete subset or the whole fauna, although it did rival Bivalvia in its ability to detect the spatial gradient (see next section). The entire class of Polychaeta is therefore an effective proxy subset of the whole fauna in southern California, but subsets within Polychaeta do not appear to be suitable on their own.

2.5.4 Bivalves are the best indicators of spatial and temporal gradients

Bivalves did not mirror the whole fauna and thus performed poorly as proxies in this study area, but out-performed both the whole fauna and other subsets in detecting the outfall gradient and its temporal decline. Bivalve richness was sensitive to decadal-scale temporal variation, but surprisingly, not sensitive to spatial variation (Figure 2.4): the temporal contrast in conditions (between Phases 3 and 1) was greater than the spatial contrast (near-

field versus far-field) even within Phase 1, implying that conditions for bivalves everywhere on today's Palos Verdes shelf permit more species than even the far-field area during peak pollution. In contrast, bivalve evenness, and composition – i.e., measures that consider numerical abundance – excelled at detecting both spatial and temporal variation on this urban shelf (Figures 2.5-2.6).

Lucinida (including *P. tenuisculpta* and other facultative chemosymbiont-bearing bivalves) was the most numerically abundant bivalve order in this dataset (77% of bivalves, 13% of whole fauna), followed by Venerida (20% of bivalves, 3% of whole fauna; part of the infaunal suspension feeding guild). However, neither clade alone could mirror patterns of the full bivalve subset or strongly detect the spatial gradient. The entire class of Bivalvia is therefore an effective indicator subset for this outfall gradient and, like the polychaetes discussed above, subsets within Bivalvia were not suitable on their own.

2.5.5 The predictive power of functional guilds for surrogacy

Changes in the relative abundance of trophic guilds has been long-appreciated as a key component of benthic response to pollution, as observed over time, along spatial gradients, and in experiments (e.g., Pearson and Rosenberg, 1978; Gray and Elliott, 2009; Riedel et al., 2012). Species that feed on plankton and other suspended material (filter feeders and photosymbiont-bearing taxa) decline in favor of those that feed on detritus and/or its microbial communities (many deposit feeders plus chemosymbiont-bearing taxa; Simonini et al., 2004; Kotta et al., 2007; Villñas et al., 2012). Many detritus-feeders are additionally tolerant of the side effects of organic enrichment, such as hypoxia, soupy mass properties, resuspended fines, and contaminants absorbed on fine particles (Maurer et al., 1999). The relative abundance of guilds that consider mobility, life habit, and trophic group is thus expected to vary significantly along environmental gradients, and perhaps more strongly so than that exhibited by taxonomic clades (e.g., Gusmao et al., 2016; Norkko et al., 2019).

Our findings on guilds as units of analysis are consistent with the general model of benthic trophic response to organic enrichment (Pearson and Rosenberg, 1978; Gray et al., 2002), with richness, evenness, and composition all differing most strongly across space during Phase 1 (high local enrichment) and weakening significantly over Phases 2 and 3.

On the Palos Verdes shelf, the power of guild diversity was comparable to either family- or ordinal-level diversity, depending on the subset. The relatively high reliability of polychaete guild diversity as a proxy is possibly tied to how closely its functional diversity resembles that of the whole fauna (S/FG = 15.8 versus 14.9); the S/FG of the less powerful bivalves and malacostracans were much lower (7.4 and 9, respectively). Because congeneric and confamilial species often encompass similar guilds (see methods), functional diversity might signal that a group will be a good proxy subset, even if genus- or family-level diversity is anomalously high.

Our observed increase in both guild richness and evenness with improved wastewater treatment (all four sets, although especially for bivalves; Figures 2.4-2.5) is consistent with the decreases in functional diversity observed elsewhere under regimes of deteriorating conditions (e.g., Gusmao et al., 2016; Norkko et al., 2019). Future studies on taxonomic surrogacy should also evaluate guilds to consolidate the utility of this often-overlooked means of analysis, which, owing to conservatism within clades, can be achieved using relatively coarse taxonomic data.

We found that the ecological disparity of guilds, rather than their number alone, should be considered, that is, the range of life habit and trophic strategies represented. Bivalves, for example, span an exceptionally broad array of guilds, from deep-burrowing chemosymbiosis to free-swimming epifauna (both suspension feeders and predators), and thus can cover a broader environmental response along a pollution or other stress gradient than the larger number but more nuanced spectrum of guilds present in polychaetes and in malacostracans. Indeed, the functional response of bivalves to improved water treatment on the Palos Verdes

shelf involves a precipitous decline in the abundance of chemosymbiont-bearing taxa in the late 1980s and the gradual increase of infaunal suspension and mixed feeders (Stull et al., 1996; LACSD, 2012; Leonard-Pingel et al., 2019). A broad range of guilds may thus be a good signal that a taxonomic subset would be an effective indicator. A means of quantifying functional disparity among taxonomic subsets would help validate this potential method of subsetting taxonomic data.

2.6 Conclusions

This case study, using an exceptionally long-term and taxonomically consistent macrobenthic dataset from the urban Palos Verdes shelf of southern California, demonstrates that coarsening the taxonomic resolution of animal identification and focusing on taxonomic subsets of the whole fauna are effective techniques, individually and in combination, to reduce the taxonomic expertise and processing efforts of monitoring in regions with a strong suspected pollution gradient. Our findings – that both polychaete and bivalve subsets detected weakening of the spatial gradient over time, thus reflecting the whole fauna, and patterns remained robust when coarsened to families or functional guilds – almost certainly reflect (a) the high strength of the original pollution gradient and (b) the preponderance of monospecific genera, low-diversity families, and phylogenetic conservatism of traits that confer guild affinities and pollution tolerances of individual species in this fauna. These results thus strengthen the positive results for taxonomic sufficiency and caveats for taxonomic subsetting that have been emerging from tests elsewhere. Insights acquired from the highly-resolved, methodologically-standardized, and broad-scoped data produced by long-term monitoring programs in well-funded regions like southern California can be used to advise the design of efficient programs for areas where the fauna is less fully documented, local taxonomic expertise and/or funds are scarce, and yet the need for environmental evaluation is still or perhaps especially urgent.

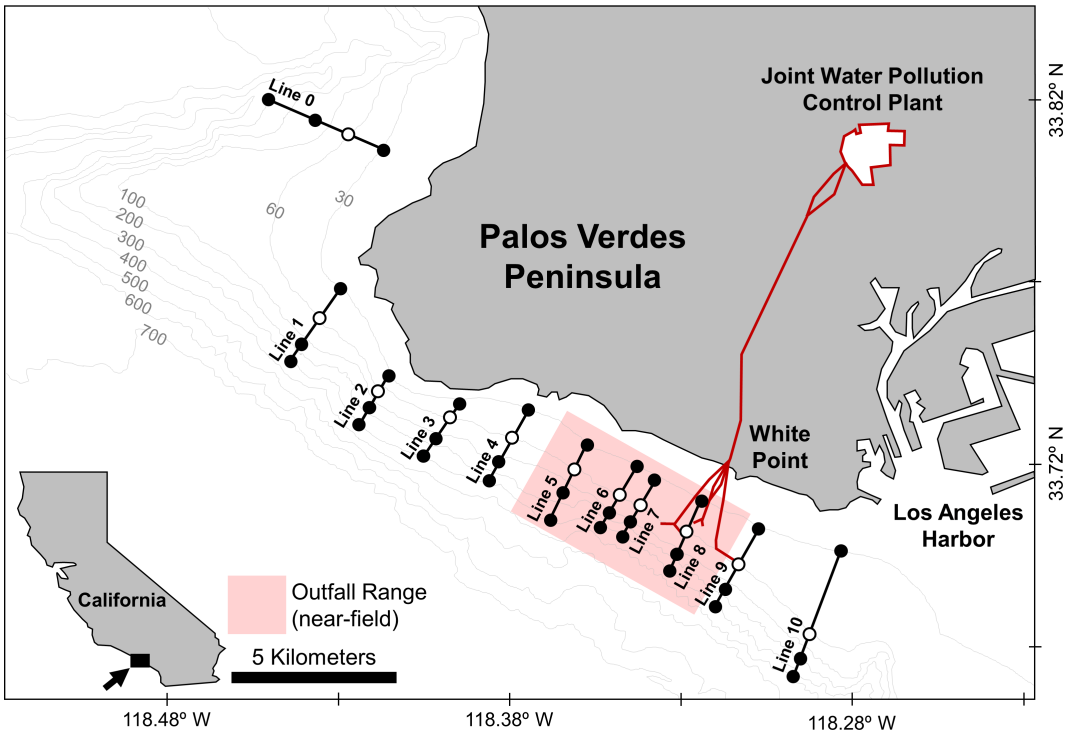

Figure 2.1: Study area off the Palos Verdes peninsula on the southern California continental shelf. Annual monitoring of macrobenthos is conducted along 11 bathymetric sampling transects (Lines 0-10) at 30, 60, 150, and 300 m depths. White dots indicate sampling stations along the 60 m isobath used in this study. Gray lines are isobaths and red lines are the outfall pipes extending from the JWPCP to White Point. Red area approximates the stations near the outfall source (near-field) defined in this study, encompassing lines 5-8, and the rest of the shelf is considered far-field (lines 0-4, 9-10). Modified from LACSD (2020).

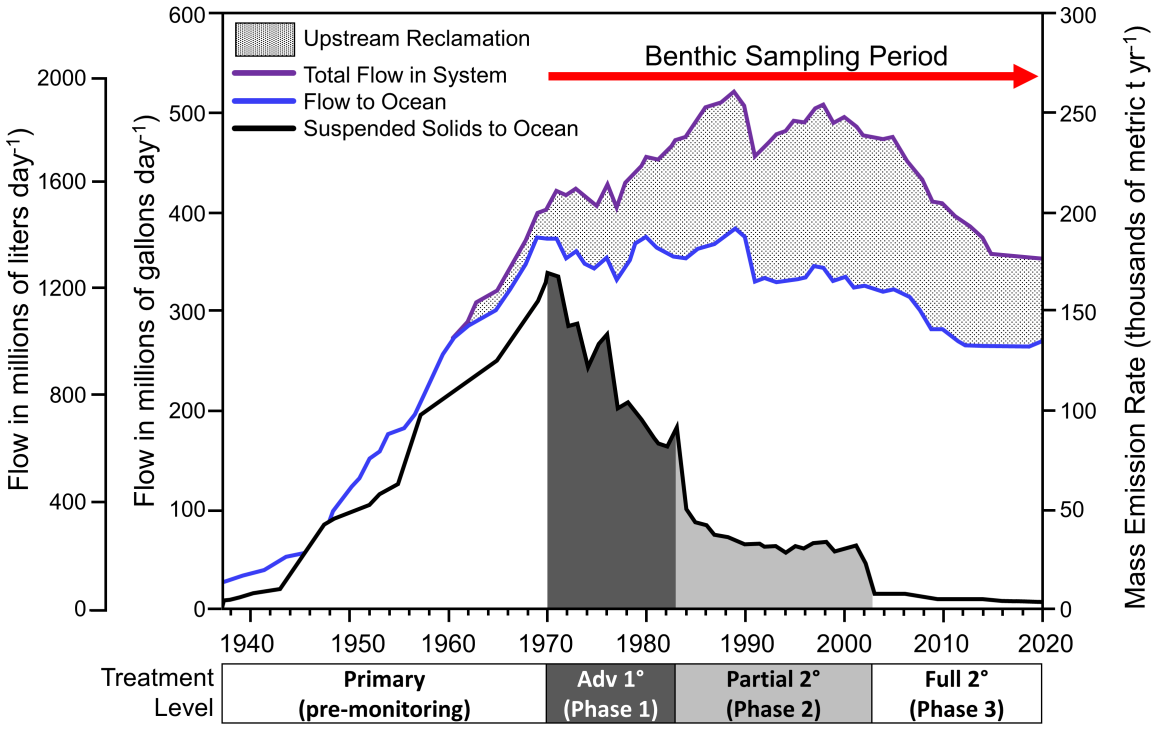

Figure 2.2: Discharge history (1 US gallon = ∼3.79 L) from the Joint Water Pollution Control Plant (JWPCP) through the White Point outfall system from 1937-2019. Suspended solid release (103 metric tons/yr) increased until the 1970s and then declined with advanced primary wastewater treatment (Phase 1), partial secondary treatment (Phase 2), and full secondary treatment (Phase 3). Biological oxygen demand also decreased strongly, by 50% over the first few years of Phase 2 and to ∼0 during the initial few years of Phase 3 (not shown; Stein and Cadien 2009). Modified from LACSD (2020).

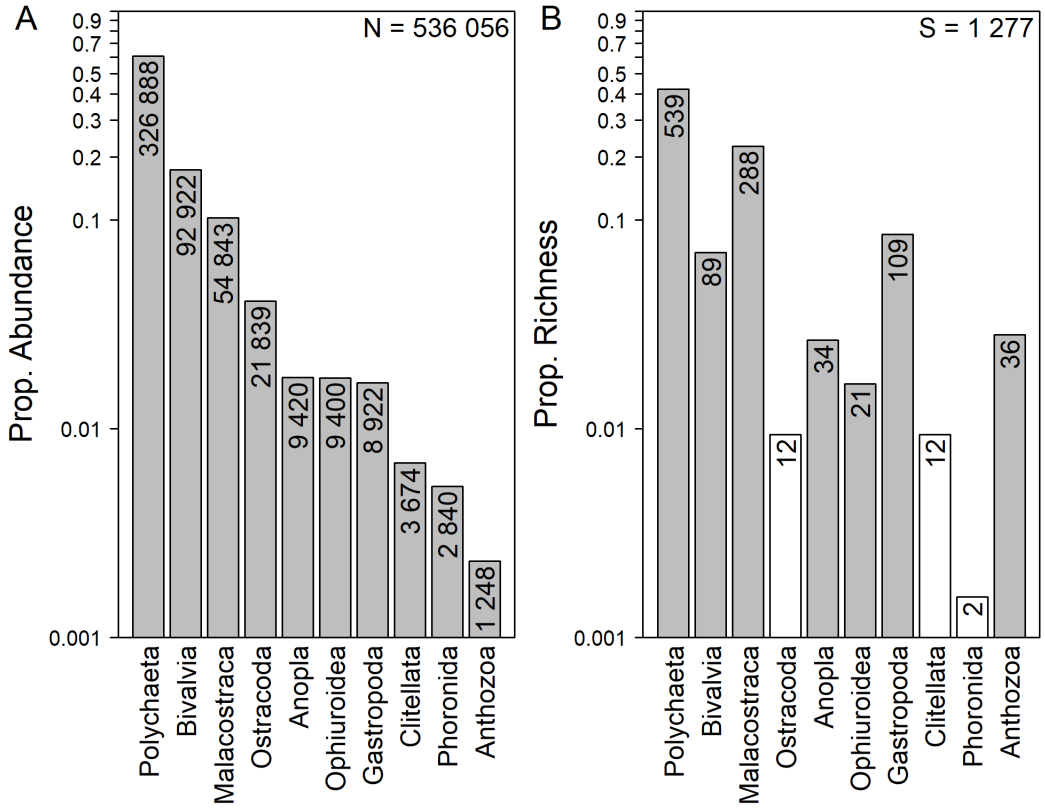

Figure 2.3: Proportional abundances (A) and species richness (B) among the ten most abundant faunal classes represented in benthic samples from 1972-2019. Numbers within bars are the raw numbers of individuals or species. The total abundance (N) and species richness (S) for the whole fauna are printed in the upper right corners. White bars in (B) indicate three classes that rank among the ten most abundant (A) but are not among the top ten in richness. Our selected taxonomic subsets (Polychaeta, Bivalvia, and Malacostraca) represent 88% of sampled individuals and 72% of sampled species during the 47 years of monitoring along the 61-m isobath of the Palos Verdes shelf.

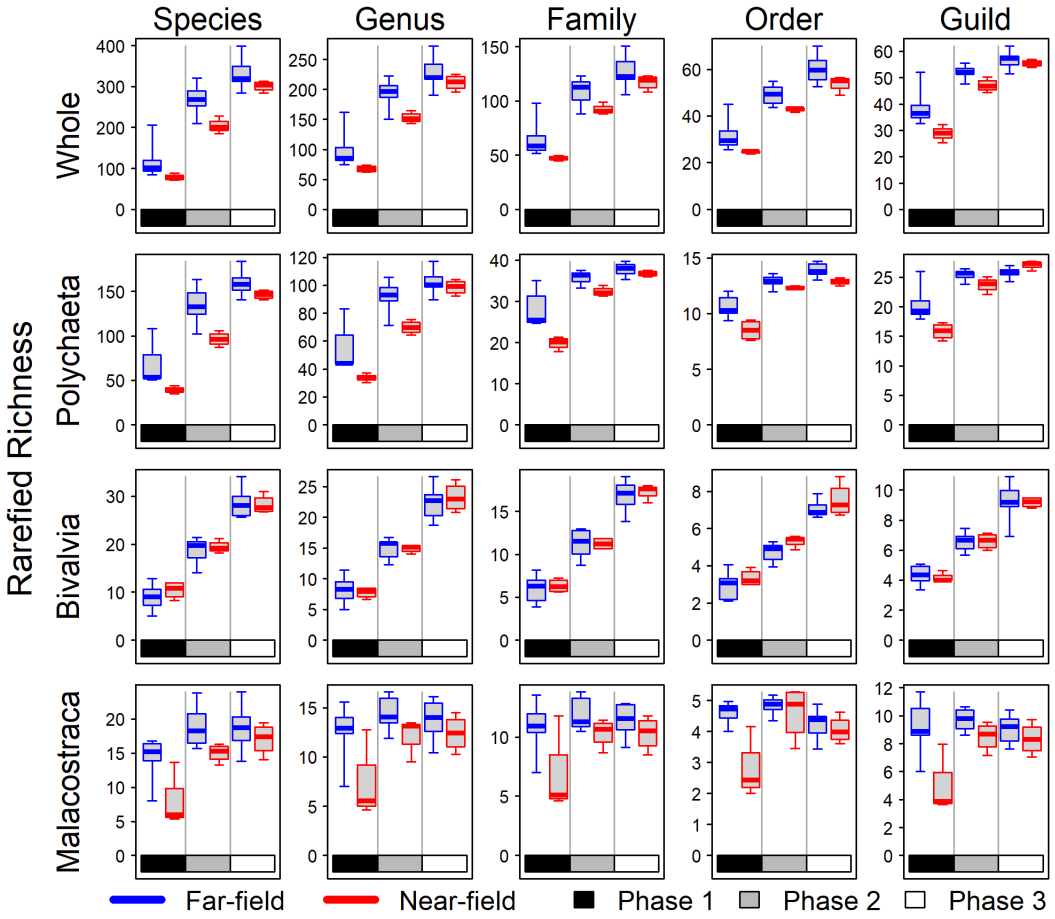

Figure 2.4: Box plots of rarefied richness for whole benthic samples and subsets (rows) and coarsened taxonomic resolutions (columns). Heavy line is median, box denotes the interquartile range (IQR), and whiskers denote entire range of values. Paired boxes compare stations at far-field (blue; seven 61-m stations) and near-field (red; four stations) areas on the shelf, as judged from sediment chemistry during the 1970s. In each graph, richness was rarefied to the smallest station-level sample size. Richness increases with improved wastewater treatment in all analyses, from Phase 1 (1974-1983) to Phase 2 (1984-2002) and Phase 3 (2003-present; shaded bars along x-axes). Polychaete patterns best mirror the whole fauna and bivalves exhibit the strongest increase across phase boundaries. For all sets except bivalves, richness is higher at far-field stations rather than near-field. This contrast is damped at coarser taxonomic levels.

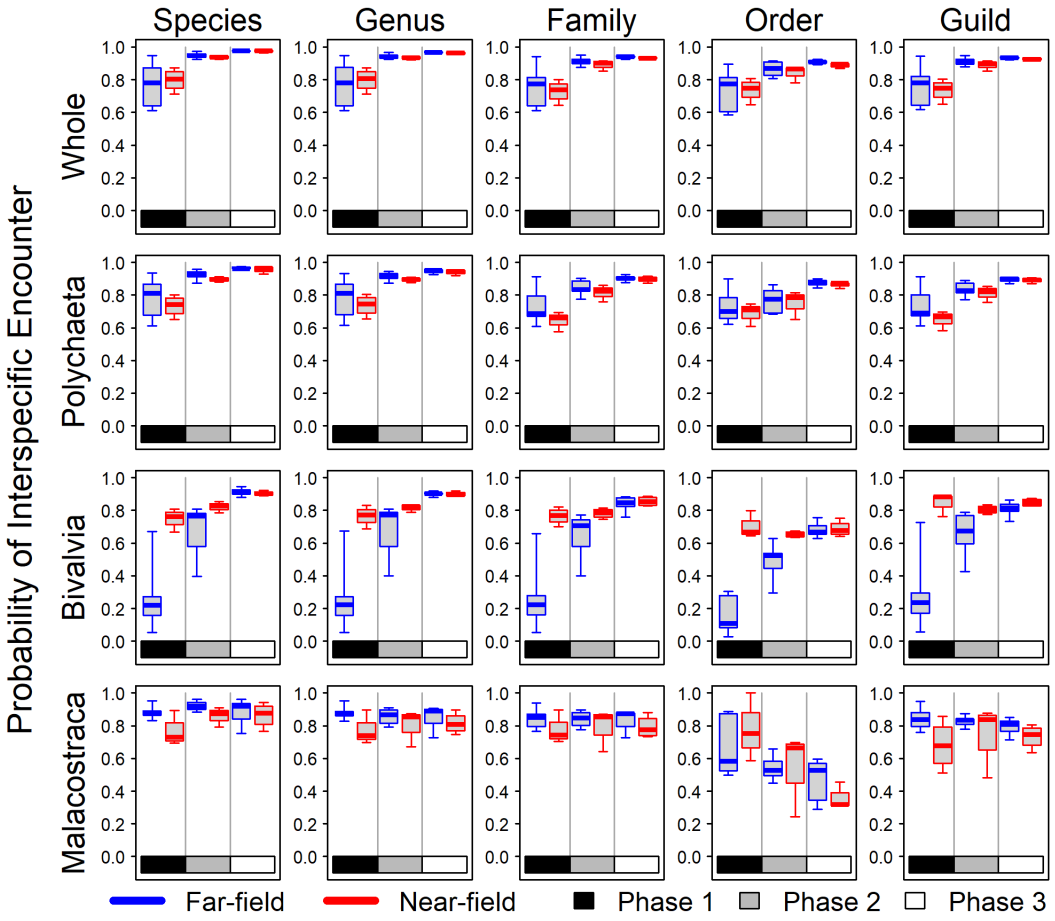

Figure 2.5: Box plots of taxonomic evenness (PIE) for whole benthic samples and subsets (rows) at increasingly coarse taxonomic resolutions (columns), with plots organized as in Figure 2.4. Evenness increases with improved wastewater treatment, but bivalves exhibit an especially strong trend arising largely from strong contrasts between near-field and far-field stations early in the treatment history. Patterns persist with taxonomic coarsening (left to right within each row) except for malacostracans at the ordinal level.

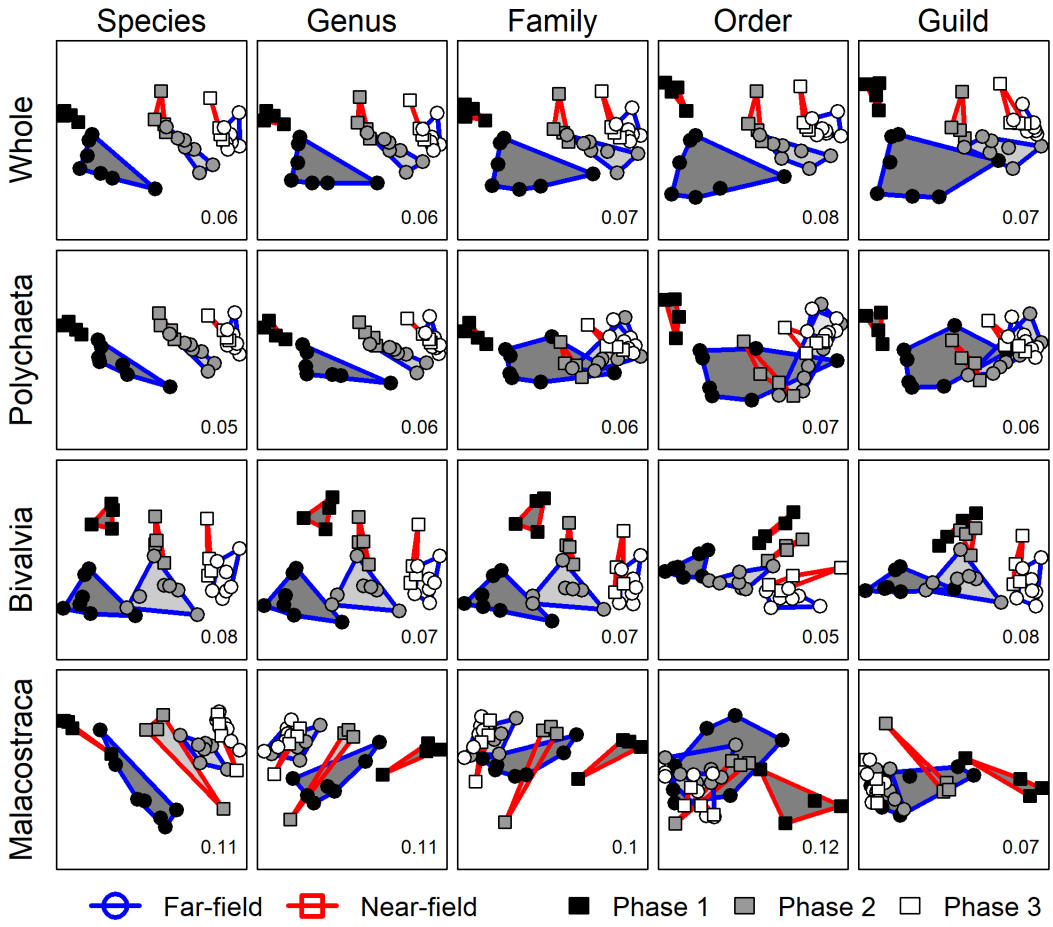

Figure 2.6: Non-metric multidimensional scaling (NMDS) plots of Bray-Curtis similarities for whole benthic samples and subsets (rows) at increasingly coarse taxonomic resolution (columns). Distances are based on square-root proportional abundances; convex hulls group samples by water treatment phase (point colors) and outfall proximity (point symbols); ordination stress values printed in lower right corner of each plot. All analyses approximately ordinate successive treatment phases along NMDS axis 1 (horizontal axis) and ordinate the spatial gradient along NMDS axis 2 (vertical axis). The taxonomic compositions of near-field and far-field stations become increasingly similar as wastewater treatment improves, both in the whole fauna (top row) and in each subset.

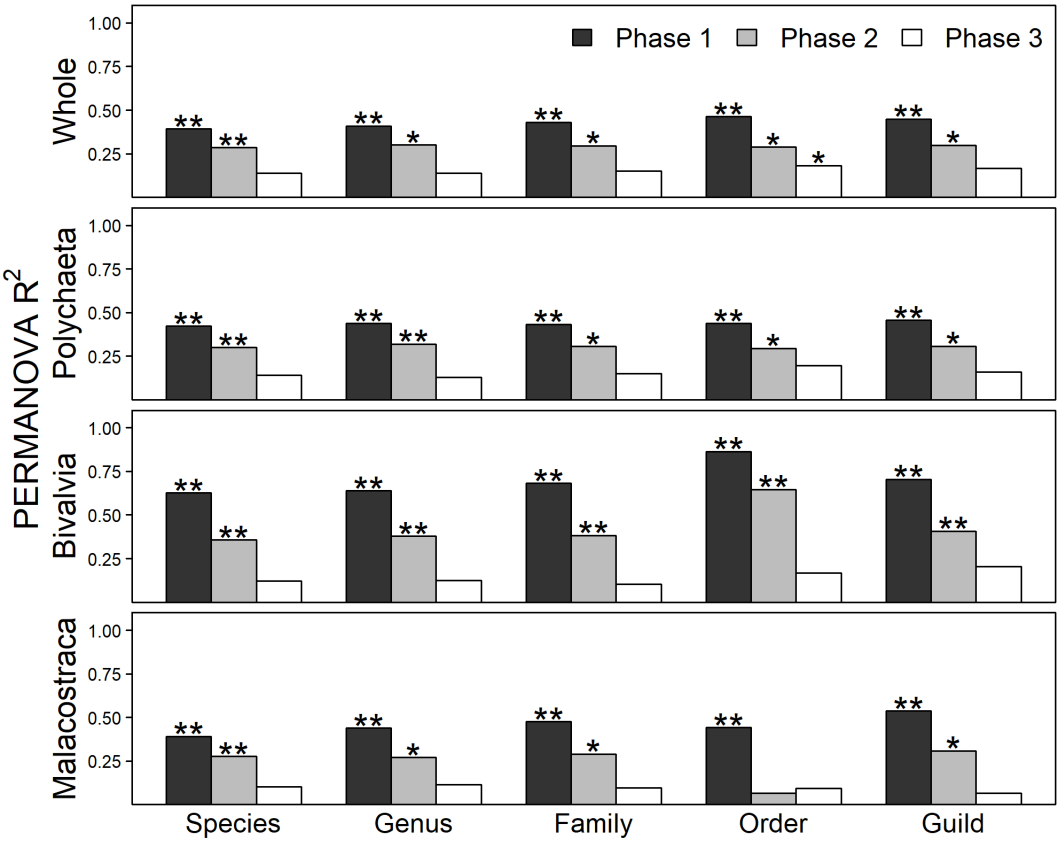

Figure 2.7: Bar plots of PERMANOVA R^2 values for compositional differences between near-field and far-field stations during each phase of wastewater treatment (bar color) for whole benthic samples and subsets (rows) at increasingly coarse taxonomic resolution (x-axis). Asterisks indicate significant p-values (* < 0.05, ** < 0.01). All taxonomic sets at all taxonomic resolutions detect spatial homogenization of community compositions with improved wastewater treatment. Polychaetes were most consistent with the whole fauna and bivalves exhibited the strongest change in R^2 values across treatment phases.

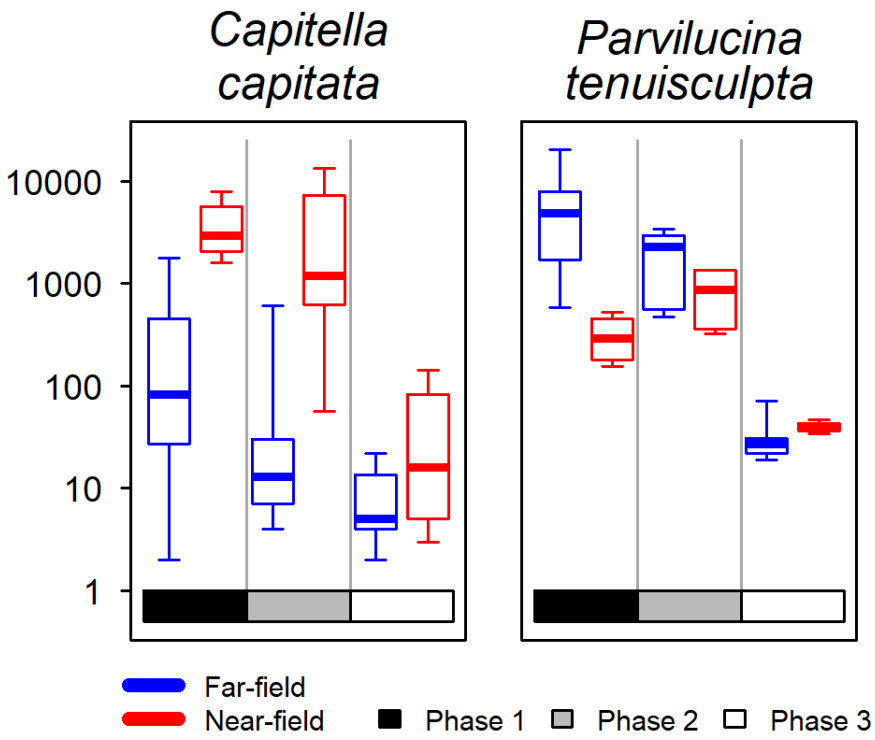

Figure 2.8: Raw abundances of the polychaete *Capitella capitata* (left) and bivalve *Parvilucina tenuisculpta* (right). Paired boxes compare near-field and far-field stations as described in Figure 2.4. *Capitella capitata* abundance is higher at near-field stations while *P. tenuisculpta* is more abundant at far-field stations. As wastewater treatment improved, the abundance of both species decreased and became more similar spatially.

Table 2.1: Common aspects of surrogacy as defined in this study.

Term	Definition
Surrogacy	Reducing the effort required to assess a biota
I. Sufficiency	Coarsening the detail/resolution/acuity of information
A. Numerical Sufficiency	Coarsening the measure of abundance (e.g., binary presence-absence or rank-abundances)
B. Taxonomic Sufficiency	Coarsening the taxonomic units (e.g., binning species into higher ranks or functional groups)
II. Subsetting	Narrowing the scope or breadth of information
A. Numerical Subsetting	Narrowing the number of specimens assessed (e.g., limit to first 100 individuals processed)
B. Taxonomic Subsetting	narrowing the taxonomic clades assessed (e.g., limit to a single class/phylum/functional group)
i. Proxy Subset	Subset that best reflects the whole fauna
ii. Indicator Subset	Subset that best detects gradient(s) of concern

Table 2.2: Representation of taxa from the whole fauna in the functional guild list by Macdonald et al. (2010). Taxa lacking a functional guild assignment were not included in the analysis of guilds. No genera represented more than one guild, but a few families represented more than one. Guilds thus represent a taxonomic resolution between the levels of genera and families.

	Richness	Taxa not assigned to a guild	Taxa representing a single guild	Taxa representing >1 guild
Species	1,277	20	1,257	
Genus	722	19	703	0
Family	347	12	302	33
Order	120	11	74	35
Class	34	2	13	19
Phylum	13	0	4	9

Table 2.3: Taxonomic richness at seven hierarchical levels (rows) and the ratio between the number of species and number of supra-specific taxa (S/T) for the whole benthic fauna and three taxonomic subsets. Higher ratios indicate higher within-taxon diversity.

	Whole		Polychaeta		Bivalvia		Malacostraca		Mean
	S	S/T	S	S/T	S	S/T	S	S/T	S/T
Species	1,277		539		89		288		
Genus	722	1.8	247	2.2	62	1.4	166	1.7	1.8
Family	347	3.7	49	11	29	3.1	89	3.2	5.3
Order	120	10.6	16	33.7	10	8.9	8	36	22.3
Class	34	37.6							
Phylum	13	98.2							
Guild	86	14.9	34	15.8	12	7.4	32	9	11.8

Table 2.4: Summary of SIMPER results for each dataset over three treatment phases (P1-P3), indicating the number of taxa that cumulatively contribute to 25%, 50%, and 75% of total Bray-Curtis dissimilarity. Parenthetical values are the proportion of the total number of taxa (S) in a dataset. Higher proportions indicate that more taxa are required to explain the same amount of dissimilarity among samples, thus dissimilarity is lower.

Cumulative	Species			Genus			Family			Order		
Contribution	P1	P2	P3	P1	P2	P3	P1	P2	P3	P1	P2	P3
Whole Fauna												
25%	6	28	54	4	17	30	3	9	15	2	4	7
50%	28	111	169	21	64	98	12	31	46	5	10	18
75%	91	288	363	67	163	214	37	78	103	14	26	39
Total S	373	975	953	279	579	723	155	282	285	121	99	106
Polychaeta												
25%	4	13	27	3	7	14	3	3	5	2	2	3
50%	16	53	82	12	27	43	7	9	13	3	4	5
75%	50	132	173	36	65	89	15	19	24	6	7	9
Total S	183	540	440	125	221	220	50	47	47	17	17	15
Bivalvia												
25%	2	4	6	1	3	5	1	2	3	1	1	2
50%	4	9	15	3	7	12	3	4	8	2	2	3
75%	8	21	29	6	14	23	4	8	14	2	3	6
Total S	40	73	61	32	52	63	19	30	26	8	9	10
Malacostraca												
25%	4	8	13	3	4	6	3	3	4	2	1	1
50%	9	25	35	7	15	19	6	9	12	3	3	3
75%	23	65	71	15	38	41	11	22	27	4	5	4
Total S	71	213	289	56	130	167	43	74	90	7	8	8

CHAPTER 3

THE POWER OF LIVING AND DEAD BIVALVES TO DETECT SPATIAL AND TEMPORAL POLLUTION GRADIENTS USING ECOLOGICAL QUALITY INDICES (BRI AND AMBI) ON THE SOUTHERN CALIFORNIA SHELF

3.1 Abstract

Benthic biologists have developed numerous index methods to score the intensity of ecological response to various human stressors including wastewater pollution, but they rarely have long time series to ascertain the decline or even the recovery of healthy conditions. Naturally time-averaged skeletal debris in the seabed ('death assemblages') might prove useful to this end: although typically dominated by very young, recently input shells, they can also include locally produced shells that are several hundreds to thousands of years old. Here, I used species-level abundance data from living benthos sampled over a 47-year time series (1972-2019) from the Palos Verdes shelf of southern California to evaluate the power of bivalves alone to (1) detect temporal and spatial gradients in pollution, (2) match the scores generated using the whole fauna, and (3) the ability of their dead-shell assemblages (sampled in 2008 and 2016) to detect the same spatial gradient, that is to retain a strong signal from the historically acute pollution of the 1960s-80s. I used two index methods: the regionally-developed Benthic Response Index (BRI) and the globally-employed ATZI's Marine Benthic Index (AMBI). Index scores generated using live-collected bivalves showed a strong positive correlation with scores based on the entire macrobenthic fauna, but gave inaccurately high values (that is, falsely indicated a need for remediation) in 20-40% of all samples, depending upon water depth and index. Adjusting the bivalve-only scores based on bootstrapped linear regressions notably improved the agreement of ecological status between bivalves and the

whole fauna. Overall, both indices were able to detect that benthic ecological quality had indeed improved between the distinctly degraded conditions of the 1970s-1980s to the arguably fully-recovered conditions of the 2000s-2010s. Death assemblages from the Palos Verdes shelf are known from independent geological studies to be significantly time-averaged, and those collected in 2016 indicated better conditions than those collected in 2008, a difference consistent with predictions (the youngest cohorts in 2016 were from better conditions than the youngest input in 2008). However, using both BRI and AMBI, death assemblages suggested that ecological quality in 2008 was as high as or higher than that registered by the living bivalve community in the 2000s, and that in 2016 the living community health had (correctly) continued to improve. The compositions of these death assemblages were notably discordant only with the compositions of living bivalves under peak pollution (1970s-80s): surprisingly, using these metrics, the death assemblages apparently carry very little signal from those geologically relatively recent, highly-polluted populations. The relatively high failure rate of living bivalves (indicating a false need to remediate), and the inability of death assemblages to detect strong improvement in ecological quality over recent past decades, are both inconsistent with findings of high power of living and dead bivalves for detecting pollution gradients using other metrics, both here and elsewhere. This weakness appears to derive from the metrics themselves: in this setting, and notwithstanding trialing both a regionally-customized and a widely used index, bivalve taxa occupy only a narrow range in the total spectrum of pollution tolerance exhibited by the whole fauna (and especially by polychaetes), a limitation that might well exist elsewhere.

3.2 Introduction

Macrobenthic communities are well known to vary strongly along natural environmental gradients and to respond rapidly to anthropogenic stressors, making them powerful bioindicators of ecological quality in marine and freshwater monitoring (e.g., Todd et al., 2019).

Declines in marine ecosystems worldwide, caused by such stressors (e.g., resource harvest, eutrophication, contamination, watershed development), of which many are associated with 20^{th} century urbanization, have increased the need for quantitative measures of ecological quality, both to identify locations needing remediation and to guide the prioritization of management resources. These tools, known collectively as biotic indices, are distinguished from conventional metrics such as richness and Shannon diversity by focusing specifically on the pollution- or other tolerance of taxa, which is used to classify samples along a semi-quantitative scale of 'Ecological Quality Status' (EQS). Monitoring agencies and government regulatory bodies around the world have developed and/or adopted various biotic indices as part of their analytical protocols, aiming to standardize the criteria by which marine habitats are assessed and managed (Rosenberg et al., 2004; Van Hoey et al., 2010; Borja et al., 2012).

However, for given regions, the oldest available data for macrobenthic community composition typically date only to the initiation of monitoring, which commonly coincides with or very close to the onset of a particular stressor, making it challenging to estimate EQS over time frames longer than a few decades (Gray and Christie, 1983) or to characterize the "before" component of a standard before-after-control-impact study design (BACI; Underwood, 1991). Without ecological information that predates ongoing and legacy stressors, sites that are used as references (the "control" component) for natural conditions might have already been shifted significantly from their historic range of variability (e.g., shifting baseline syndrome; Pauly, 1995; Villnäs and Norkko, 2011), a situation recognized by many agencies that refer to such sites as references rather than controls. Using data from less-than-fully-natural reference sites could lead both to underestimating the magnitude of ecological strain imposed by long-term stressors and to defining improper target areas for conservation, mitigation, and restoration (Borja et al., 2012; Atmore et al., 2021). Retrospective insights into pre-monitoring and pre-stressor conditions are thus needed to supplement the high-quality data generated by even long-term monitoring programs.

Inferences from the composition of naturally accumulated macrobenthic 'death assemblages' are a promising means of acquiring such critical data. Death assemblages are the skeletal remains of previously living organisms (e.g., bones, shells, sclerites, etc.) that accumulate naturally in the surface mixed layer of the seabed and, owing to their postmortem durability, temporally aggregate past generations. Although typically dominated by specimens contributed during the most recent few decades to centuries, depending upon sediment accumulation rates and biogenic mixing, these assemblages from the upper mixed layer of the seabed also include a long tail of specimens that are 10 to 100x older in tropical to warm-temperate settings (see review by Kidwell, 2013; also Nawrot et al., 2022). This legacy of skeletal material from past generations can be used as a time-averaged summary of long-term biodiversity. Previous studies, focused on mollusks in level-bottom seabeds, have documented that shells are overwhelmingly locally generated (i.e., are not moved out-of-habitat; e.g., Kowalewski et al., 1998; Hyman et al., 2019) and that the effects of age-mixing on species richness, evenness, beta diversity, and composition are predictable (see summary in Kidwell and Tomašových, 2013). Death assemblages, by virtue of being acquired concurrently with living infaunal communities, are a uniquely inexpensive and readily obtainable complement to biomonitoring. They simply need to be retained rather than discarded with other sieve residue, and then be processed for counterpart 'dead data' on species' abundances. Death assemblages of shelled mollusks have proven to have high compositional agreement with co-occurring living bivalve communities ('living assemblages') in relatively pristine and/or steady-state coastal areas, based on early studies (e.g., Warme, 1969), meta-analysis and modeling (Kidwell, 2001; Tomašových and Kidwell, 2009, 2013), and subsequent focused tests (e.g., Hyman et al., 2019). Importantly, meta-analysis (Kidwell, 2007) demonstrated that subtidal habitats exhibiting poor live-dead agreement are consistently associated with areas having a known history of anthropogenic disturbance, in particular eutrophication. Live-dead discordance as a signal of anthropogenic shifting of long-term baselines has been

corroborated by many field studies of estuaries and open shelves, demonstrating its ability to recognize eutrophication (Korpanty and Kelley, 2014; Gilad et al., 2018; Tweitmann and Dietl, 2018), invasive species (Yanes, 2012; Chiba and Sato, 2013; Steger et al., 2022; Kokesh and Stemann, 2023), bottom trawling (Kidwell, 2009), solid-sediment pollution from historic land use (Tomašových and Kidwell, 2017; Sander et al., 2021), and climate change (Powell et al., 2017; Meadows et al., 2019). As part of a new interdisciplinary field called conservation paleobiology, an increasing number of geologists and paleontologists are now applying many kinds of geohistorical data including from death assemblages to inform the management of modern biodiversity (Dietl and Flessa, 2011; Dietl et al., 2015; Dillon et al., 2022).

The ability of death assemblages to capture evidence of past pollution via more specialized EQS metrics is, however, less clear. For example, re-analyses of published datasets to assess the ability of shelled-mollusks to replicate EQS gradients in the whole fauna found that this subset can be quite powerful (e.g., Dietl et al., 2016; Pruden et al., 2021, and see many studies finding that bivalves have high power to detect spatial variation in more conventional diversity measures, e.g., Włodarska-Kowalczuk and Kędra, 2007; Bevaliqua et al., 2009; Kokesh et al., 2022b). However, death assemblages associated with the data used to quantify those gradients were not evaluated, and thus the persistence of EQS signals through the gauntlet of postmortem destruction and time-averaging remains uncertain. Here, I evaluate the ability of bivalves alone – both alive and dead – to serve as proxies for the EQS calculated using the whole microbenthic fauna on a continental shelf having an exceptionally long-term database on macrobenthic communities and environmental conditions related to point-sourced wastewater emissions. Annual sediment sampling of the Palos Verdes shelf, southern California, mandated by the US Clean Water Act, has produced a taxonomically consistent, quantitative species-level dataset on macrobenthic abundance for 44 stations since 1972, covering the entire period of decreasing emissions (Stein and Cadien, 2009; LACSD, 2022). I estimated the EQS for the whole fauna and bivalves only from the live-collected database, and

for bivalve death assemblages saved from the 2008 and 2016 surveys, using two well-known biotic indices: (1) the Benthic Response Index (BRI; Smith et al., 2001; 2003) developed by agencies in Southern California and widely used in local monitoring (e.g., City of San Diego, 2022; LACSD, 2022; OCSD, 2023), and (2) the ATZI's Marine Benthic Index (AMBI; Borja et al., 2000) developed originally for European waters, but since applied globally. My analytical framework consisted of two phases: (1) testing the power of the bivalve subset as a surrogate for the whole fauna (i.e., a test of taxonomic surrogacy; Moreno et al, 2007; Kokesh et al., 2022b), and (2) comparing the EQS index scores of living bivalves to those of the co-sampled but naturally time-averaged bivalve death assemblages.

3.3 Methods

3.3.1 Study area

The Palos Verdes shelf off of Los Angeles County is a narrow, ∼25 km long platform with a steep slope to the basin floor at 700 m (Figure 3.1). Warm waters from Baja California, Mexico, are transported northward by the Southern California Countercurrent before dissipating at Point Conception, which marks the northern edge of the Southern California Bight. Wastewater pollution on the Palos Verdes shelf began with the installation in 1937 of a 34 m deep outfall pipe at White Point. Two additional outfall pipes were installed in 1956 and 1966 at 60 m depth and remain in use today. The total suspended solids emitted increased over the course of mid-20^{th} century urbanization, peaking around 1970 at 150,000 metric-tons per year (Figure 3.2). Wastewater was subject to primary treatment until passage of the US Clean Water Act in 1972. Solid emissions dropped sharply with the implementation of advanced primary treatment in the 1970s, partial secondary treatment in the 1980s-1990s, and full secondary treatment in the early 2000s (Figure 3.2). DDT, PCBs, metals, and other contaminants in the effluent also dropped sharply with improved treatment, as did organ-

ics and biological oxygen demand, permitting the onset of biological recovery in the 1970s (Swartz et al., 1991; Stull et al., 1996; Lee et al., 2002; Stein and Cadien, 2009). This strong point-sourced spatial gradient in sediment contamination extended outward in all directions from the outfall, but was mostly deflected northwestward from the outfall area by the Southern California Countercurrent. Today, surface sediments along this historic, multi-decadal pollution gradient are chemically indistinguishable, and most of the legacy pollutants are buried beneath the sediment-water interface (Stein and Cadien, 2009; LACSD, 2022).

The Los Angeles County Sanitation Districts (LACSD) began annual monitoring of the macroinvertebrate benthic communities in 1970. The original sampling array consisted of 10 southwest-facing bathymetric transects, each with four stations at 30, 61, 152, and 305 m water depths (Lines 1 through 10; Figure 3.1). Only data from the 30, 61, and 152 m depth stations are used here because the 305 m stations usually yielded too few bivalve dead shells for analysis. The 61 m station on Line 8 approximates the outfall source. Line 0 was added in 1974 on the northern edge of the Peninsula into the head of Redondo Canyon as a reference transect that was sufficiently down-current to have minimal pollution stress. This ∼50 years of annual monitoring has enabled detection of long-term changes in benthic community diversity and composition concomitant with improved wastewater treatment, including, by the 1990s, a more homogeneous macrobenthos along isobaths (Stein and Cadien, 2009; Kokesh et al., 2022b).

3.3.2 Processing of living and death assemblages

The LACSD living benthic invertebrate database includes all sampling records since 1972, and we use this full database up to 2019. Stations were sampled semi-annually in February and August until 2006, when the agency moved to annual summer sampling. Here, I used only the summer samples from all years. Samples were collected using Shipek grabs until 1980, when the agency switched to 0.1 m2 van Veen grabs. Three replicate Shipek grabs

from pre-1980 samples were pooled to match the sediment volume of one van Veen grab. For macrofauna, sediments were sieved by the agency through a 1-mm mesh with seawater onboard research vessels, immediately fixed in 10% formalin, and within a few months shifted to ethanol for storage. Taxonomists at LACSD sorted all invertebrates from samples, identifying organisms to the finest resolution possible (i.e., species). The taxonomy used by LACSD, other regional municipal agencies, and the Southern California Bight Regional Monitoring Program has been standardized and updated regularly by the Southern California Association of Marine Invertebrate Taxonomists (SCAMIT, 2021). All species assignments are thus held to a single, internally-consistent taxonomy, which was also used for identifying dead shells. In addition to the whole macrobenthic fauna, I extracted counts for bivalve taxa alone to generate a faunal subset matrix.

Sediment residue from van Veen grab samples is usually discarded by monitoring agencies after removal of all live-collected macroinvertebrates and completion of a quality assurance process. Here, samples from their 2008 and 2016 surveys were instead saved, decanted of ethanol, and picked of dead bivalve shells from the air-dried sediments. Any shell that retained at least half of an identifiable hinge was counted as an individual and taxonomically identified to the level of species when possible. These counts were not halved to compensate for each living individual having two valves because no other mollusk classes were being considered (contra the general concerns of Kowalewski and Hoffmeister, 2003). The wide range of body sizes and shell conditions among the individuals of a given species also argued that each valve represented a unique individual (following Gilinsky and Benninton, 1994; Kidwell, 2009). Death assemblages from the 2016 survey were not completely picked of all identifiable organisms, only for the most abundant taxa, which likely led to underestimating species richness. However, the biotic indices used in this study (see next section) are sample-size independent, and thus the proportional abundances of organisms processed were still able to be used to estimate EQS.

3.3.3 Biotic indices

The Benthic Response Index (BRI; Smith et al., 2001; 2003) is a function of the abundance of taxa weighted by their relative tolerance of anthropogenic stressors. It was developed collaboratively by multiple agencies in southern California as an intuitive metric of the ecological condition of a given benthic macroinvertebrate community. The score for a given sample is calculated as:

$$BRI = \frac{\sum_{i=1}^{S} a_i^f p_i}{\sum_{i=1}^{S} a_i^f} \tag{3.1}$$

where s is the total number of species in the sample that have been assigned some pollution tolerance value, a_i is the numerical (or proportional) abundance of the i^{th} species, p_i is the pollution tolerance value of the i^{th} species, and f is an exponent that transforms the abundance values. Following Smith et al. (2001), I set $f = 1/3$ to apply a third-root transformation. The pollution tolerance score for a given species – and whether it's even assigned one – varies depending on the water depth from which the sample was taken (e.g., shallow 10-35 m, middle 25-130 m, or deep 110-324 m): three different lists of pollution tolerance values are thus available for coastal shelf settings. I used the shallow-depth species list to assign species from samples taken along the 30 m isobath, the middle depth list for samples from the 61 m isobath, and the deep depth list for samples from the 152 m isobath. The pollution tolerance values used here were generated by Smith et al. (2001) based on an ordination of species occurrences with independent environmental data, mostly sediment chemistry. The lists have since then been modified by LACSD to account for new taxonomic identities. Taxa that did not appear in the appropriate depth list were not included in the BRI analysis.

BRI scores have no lower or upper bounds but typically range between zero and 100. Smith et al. (2001) defined threshold scores to categorize samples as falling into one of five EQS's (Table 3.1). Following Gillett et al. (2017), I simplified my interpretations of EQS

to a binary system: 'Good' conditions (i.e., requires no remediation) had index scores <34, whereas 'Poor' conditions (i.e., samples that require some level of remediation) had index scores ≥ 34.

ATZI's Marine Benthic Index (AMBI; Borja et al., 2000) was one of several index methods developed for use in European waters but it has since become a globally-popular metric due to its ease of implementation. Like BRI, the AMBI procedure calculates a single numerical score for a sample based on abundance-weighted tolerance values that have been assigned a priori to certain species, using expert knowledge. However, rather than each taxon being assigned a unique tolerance value, AMBI categorizes taxa into one of five Ecological Groups (EGs) labeled by Roman numerals: taxa assigned to EGI are classified as the most sensitive to disturbance while those assigned to EGV consist of first-order opportunists, i.e. taxa that tend to be the first colonists of azoic sediments. The score for a given sample is calculated as:

$$AMBI = \sum_{i=1}^{5} 1.5(i-1)g_i \tag{3.2}$$

where g_i is the proportional abundance of all taxa that have been assigned to the i^{th} Ecological Group (EG; I, II, III, IV, or V). AZTI provides a regularly-updated global list of benthic marine taxa and their EG assignments as part of a user-friendly software tool for calculating AMBI (http://ambi.azti.es/). For this study, I employed a modified species list generated by taxonomic experts for use in US waters (Gillett et al., 2015; Pelletier et al., 2018). If a particular species in the LACSD dataset was not included in the list, I either assigned it to the same EG as its genus (if that was on the list), or did not include the species in the AMBI analysis. AMBI scores are bound to a range from zero to six, with higher values indicating increasingly-strained samples. Like BRI, four thresholds are defined to classify samples along an ordinal scale of EQS (Table 3.1). However, similarly to BRI, I simplified interpretations to a binary system: 'Good' conditions were signified by samples with AMBI

scores <3.3 and 'Poor' conditions by samples with AMBI scores ≥3.3 (Table 3.1).

3.3.4 Data analysis

All data analyses were performed using R version 4.2.3 (R Core Team 2023). Due to the patchy distribution and overall sparsity of dead shells along the 305 m isobath, I excluded those stations from the analysis. Numerical abundance, species richness, BRI, and AMBI were calculated for annually-sampled benthic community data from the remaining 33 stations for both the whole fauna and the bivalve subset. Corresponding index scores generated from the whole faunas ('whole-BRI' and 'whole-AMBI') and their bivalve subsets (henceforth, 'bivalve-BRI' and 'bivalve-AMBI') were plotted for each of the three isobaths (30 m, 61 m, and 152 m; Figure 3.1). Points that plotted above the one-to-one line represented cases where bivalve-generated scores underestimated true whole-fauna scores (i.e., conditions appear better than they should), while those that plotted under the one-to-one line represented cases where bivalve-generated scores overestimate true scores (i.e., conditions appear worse than they are). For each plotted relationship, I calculated the linear regression of the full model to generate a line-of-best-fit, with bivalve-generated scores acting as the independent variable for which to estimate the whole fauna's score as the dependent variable. To test for independence of the resulting model, I bootstrapped 1,000 resampling-with-replacement events to calculate new regressions from which mean linear coefficients and their 95% confidence intervals could be produced. These mean linear coefficients were used to adjust the directly-calculated raw bivalve scores.

BRI and AMBI scores generated with the whole fauna and bivalves (raw and adjusted via regression) were assigned to either the 'Good' or 'Poor' EQS as defined in Table 3.1. If the bivalve-generated score of a sample resulted in the same 'Good' or 'Poor' EQS as its corresponding whole-fauna score, the sample was labeled as indicating the 'Same Good/Poor' status. If bivalve-generated scores resulted in a different EQS, the sample was flagged as

either indicating a 'Better' status (i.e., bivalves indicated 'Good' when the whole fauna indicated 'Poor') or 'Worse' status (i.e., bivalves indicated 'Poor' when the whole fauna indicated 'Good'). The proportional frequencies of 'Same Good', 'Same Poor', 'Better', and 'Worse' EQS assignments were assessed for each of the three isobaths both separately and combined.

BRI and AMBI scores were calculated for bivalve death assemblages rescued from sets of samples collected in 2008 and in 2016, and their scores were adjusted using the same linear regressions determined for the living fauna.

Boxplots were produced to assess decadal changes in mean EQS using the scores generated from the annually-sampled living whole fauna, living bivalve subset, and death assemblages. Finally, I evaluated spatial patterns in BRI and AMBI scores along isobaths and with depth across the Palos Verdes shelf by plotting the mean score for each station during each decade. These scores were plotted along an x-axis representing an isobath from Line 0 to Line 10 along the shelf (Figure 3.1). The 60-m isobath along Line 8 was flagged as the location most proximal to outfall and thus pollution point source.

3.4 Results

3.4.1 Diversity and density of living and death assemblages

Mean numerical abundance (density, i.e., living animals per van Veen sample) and species richness of the whole fauna varied across both space and time (Table 3.2). Spatially, animal densities were highest along the shallow (30 m) and middle (61 m) isobaths, but richness was highest along the shallow isobath. Temporally, mean densities were highest in the 1970s-1980s and generally declined over time. In contrast, species richness increased notably over the decades and across the entire shelf, peaking in the 2000s-2010s. The decrease in density and increase in richness signals an overall increase in community evenness over time, i.e., in

the equitable distribution of individuals among taxa (decreased dominance).

For live-collected bivalves, mean density declined over the decades – a three-fold decline in shallow waters and 10-fold decline in deep waters – while mean richness increased very slightly (by only a few species; Table 3.2). Densities were highest along the deep isobath during the 1970s-1980s. The proportion of the density and richness of the whole fauna represented by bivalves increased consistently with water depth and declined consistently over the decades. Bivalves represented 14% of whole-fauna density in the 1970s but only 5% in the 2010s in shallow water, dropped from 36 to 4% in middle water depths, and dropped from 68 to 12% in deep water.

Bivalve death assemblages collected in 2008 also exhibited a depth gradient, but it was the opposite of that for living bivalves: both mean abundance and richness of death assemblages decreased with water depth. Death assemblages collected in 2016 exhibited a strong increase in density and a slight increase in richness from middle to deep waters (the only isobaths with data; note that these samples were not picked for all rare species, and thus the reported sample richness is under-determined).

3.4.2 Correlation between whole-fauna and bivalve-only index scores

Both the BRI and AMBI index scores for the whole fauna were significantly and positively correlated with counterpart scores calculated using only the bivalve subset of the community, and this was true for data from all three isobaths (Figure 3.3, Table 3.3). These correlations were not, however, consistently 1:1 (dashed reference line in each plot) or, at the least were offset from it along the intercept. For BRI, the slopes were quite close to 1:1 (ranged 0.71 to 1.14) but had notably lower intercepts, which became ever-lower with water depth. Bivalve-BRI thus tended to overestimate the whole-BRI, that is suggest that local conditions were worse than indicated by the whole fauna, with many samples plotting in the lower-right quadrant (Figure 3.3). For AMBI, slopes were consistently shallower and diverge more

strongly (range 0.42 to 0.64), and the intercepts were more variable with no clear depth gradient (Table 3.3).

The slope and intercept coefficients for all full models were captured by the 95% confidence intervals of the bootstrapped models and were nearly identical to the mean bootstrapped coefficients (Table 3.3). Mean bootstrapped coefficients were nonetheless used to derive all adjusted bivalve-BRI and bivalve-AMBI scores (displayed in Figure 3.4). Notably, the adjusted bivalve-BRI scores dramatically increased the percentage of samples where bivalve-BRI and whole-fauna-BRI agreed. For example, using raw (directly-calculated) bivalve-BRI scores, bivalves produced the same EQS for 841 (59%) if the total 1,436 samples: 47% of all samples classified both their whole-fauna and bivalve-only composition as indicating conditions were 'Poor' and 12% classified the sample consistently as 'Good' (far-right column in top-left panel; Figure 3.5). The percentage of samples where both the whole-fauna and the bivalves-only yielded a 'Poor' EQS was consistent among isobaths, but the percentage of samples indicating the same 'Good' EQS declined with depth (28% to 2%), whereas the percentage of samples where the bivalves classification was 'Worse' than that produced by the whole fauna increased (23% to 53%). The counterpart agreement of bivalves and whole-fauna using adjusted bivalve-BRI was 41% of samples agreeing on 'Good' and 42% agreeing on 'Poor' , with a combined frequency of 83% agreement, and no apparent variation with water depth (far-right column in top-right panel; Figure 3.5).

In contrast, using raw bivalve-AMBI scores, a high proportion of samples (83%) reported the same EQS as the whole-AMBI, with the vast majority of these agreeing that the shared EQS was 'Good' (far-right column in bottom-left panel; Figure 3.5). Using adjusted bivalve-AMBI scores, only a few additional samples indicating the same EQS as whole-AMBI (86%; far-right column, bottom-right panel; Figure 3.5). The proportion of samples misclassified as 'Worse' status by the adjusted bivalve-AMBI was slightly higher in shallow water while the proportion better, while the proportion of samples misclassified as 'Better' status increased

only slightly.

Adjusting bivalve-BRI scores thus notably improved their ability to indicate the same EQS as the whole fauna, but the power of the same adjustment procedure applied to bivalve-AMBI scores was much less.

3.4.3 Spatio-temporal variation in BRI between whole fauna, living bivalves, and dead bivalves

Median BRI scores generated by both the whole fauna and the living bivalves demonstrated similar stepwise declines over the last five decades, indicating improved benthic conditions (Figure 3.6). For each decade, however, raw bivalve-BRI scores tended to overestimate the whole-BRI scores, mistakenly indicating samples from the 1990s-2010s as requiring remediation. Adjusted bivalve-BRI scores rectified this discrepancy (third row; Figure 3.6), with median scores consistently falling on the same side of the threshold value of 34 as did BRI scores calculated using the whole fauna.

Bivalve death assemblages (white and dark boxes; Figure 3.6) from all isobaths yielded median raw and adjusted BRI scores that were very similar to those of living bivalves during the 2000s-2010s, contrary to expectations. The IQRs of samples from the 2008 death assemblages were consistently narrower than their living-bivalve counterparts; they were positioned lower than the counterpart living bivalves of the 2000s in shallow and middle water depths (i.e., indicating better conditions than counterpart living bivalves) but matched the living bivalves in deep water. Data from the 2016 death assemblages, which were available only for middle and deep water depths and excluded information on rare species, had median raw and adjusted BRI scores that, as in 2008 death assemblages, were either lower than the living bivalves from their decade (middle water depth 2010s) or similar to counterpart living bivalves (deep water).

Considering median BRI scores at each monitoring station along the Palos Verdes shelf

(Figure 3.7), BRI scores based on the whole fauna were highest in the vicinity of the wastewater outfall (61 m at Line 8, gray bar) and declined to lower BRI scores with distance both up-current (east, right) and especially down-current (west, left); this spatial gradient was strongest during the 1970s-1980s and in middle and deep waters. This spatial gradient faded over subsequent decades, with only a faint signal retained along the deep isobath in the 2010s. Raw and adjusted bivalve-BRI from the 1970s-1980s revealed a similar spatial gradient in middle water depths, including high BRI values extending westward (left) from the Line 8 outfall. However, bivalves exhibited less spatial variation than the whole fauna along the deep isobath, and no spatial signal along the shallow isobath (middle and lower rows in Figure 3.7). Bivalve death assemblages from 2008 – the white circles in Figure 3.7, with more complete data than the 2016 assemblages – also exhibited (1) no spatial gradient along the shallow isobath (like living bivalves from the 1980s onward), (2) a slight gradient along the deep isobath consistent with the living 1970s scores, and (3) a small peak near Line 8 along the middle isobath, also consistent with the living 1970s gradient.

3.4.4 Spatio-temporal variation in AMBI between whole fauna, living bivalves, and dead bivalves

AMBI scores demonstrated similar declines in median values over the last five decades using both the whole fauna and living bivalves (colored boxes in Figure 3.8). Scores typically decreased notably from the 1970s to the 1980s and/or the 1990s, and were then stable for the 2000s and 2010s, when IQRs were also much reduced. The bivalves did not yield medians notably better or work than the whole fauna, in contrast to when using BRI scores. Adjusted bivalve-AMBI reduced the IQRs and tended to decrease medians slightly, shifting most of the point distribution into the 'Good' EQS range.

AMBI scores from bivalve death assemblages (white and dark boxes; Figure 3.8) yielded consistently low ('Good') scores. These scores were, in shallow water, comparable to those

generated from living bivalves in the 2000s-2010s, whereas dead-AMBI scores in middle and deep water were slightly higher (poorer EQS) than those of counterpart living bivalves, that is, were shifted slightly toward living bivalve AMBI scores from the 1970s-80s (as expected for time-averaged death assemblages). Scores from the 2016 death assemblages were, as with BRI, detectably lower (better conditions) than those from the 2008 assemblages.

Median AMBI scores for the whole fauna at each monitoring station along the Palos Verdes shelf (Figure 3.9) yielded a distinct gradient similar to that produced using BRI, with higher scores near the outfall and lower scores with distance both up-current (east, right) and, especially, down-current. However, this spatial gradient was apparent only in middle water depths and even there flattened by the 1990s, in contrast to the spatially broader and temporally more persistent gradients evident using BRI on the whole fauna (Figure 3.7). Along the other isobaths, whole-fauna AMBI was laterally variable (shallow water) or exhibited only a broad low arch in values (deep water; Figure 3.9). Bivalve-AMBI scores, both using the raw and, to a less extent the adjusted values, exhibited a variety of patterns: highest values (poorest) in most distal down-current stations along the shallow isobath (near Line 2), highest values near both the up-current (Line 10) and down-current ends of the middle isobath, and a similar but damped pattern along the deep isobath. All of these settled to a negligible gradient of consistently low values ('Good' EQS) by the 1990s or 2000s. Bivalve death assemblages also exhibited a different spatial gradient along each isobath: consistently low or moderate scores along the shallow and deep isobaths, respectively, and, along the middle isobath, far down-current of the outfall, at Line 2, as also observed in living bivalve AMBI scores.

3.5 Discussion

Although community-based biotic indices were designed and are generally applied to data on the whole macrobenthic fauna, I found that BRI and AMBI scores generated using only the

bivalves are able to detect the same spatial (along-shelf) and temporal variability detected using the whole fauna (Stein and Cadien, 2009; LACSD, 2022), namely increasing ecological quality with distance from the outfall and over a 50-year history of improved wastewater treatment. However, depending on the index method, these bivalve-generated scores tended to either overestimate or underestimate the EQS based on the entire fauna (Figure 3.3), which could potentially impact management decisions. Here, because I had whole-fauna data for the entire spatiotemporal matrix, I could evaluate the magnitude of offset of taxonomic subsets from the whole fauna, allowing me to transform them to better approximate whole-fauna EQS (Table 3.3; Figure 3.4). Indeed, adjusting raw bivalve-BRI scores resolved the tendency for score overestimation and improved their accuracy in identifying the same EQS as would result from using the whole fauna (Figure 3.5). While these linear regression coefficients may serve as a useful transformation of bivalve-only surveys along the Palos Verdes shelf, it must be emphasized that they may not adequately represent the relationship between bivalve- and whole-BRI elsewhere in southern California without further cross-validation (see below).

The two dead-shell assemblages assessed here, sampled eight years apart, indicated that a signal of the improving quality of the living benthos was indeed becoming incorporated into the time-averaged surficial sediment record: the 2016 death assemblages consistently yielded lower scores than the 2008 death assemblages (Figures 3.6-3.9). This change in death assemblage composition is in the predicted direction, and it is interesting how distinct it is given that the two death assemblages differ in age by only 8 years and that the density of living bivalves during this interval (and thus their input of newly dead shells) is quite low (Table 3.2). With respect to the expectation that time averaging would have incorporated many shells produced during the highly-disturbed mid-20^{th} century into the dead-shell record, BRI scores generated by death assemblages were surprisingly low, and AMBI scores only fitted this expectation slightly better.

3.5.1 Caveats and considerations of index selection

Owing to their differing mathematical procedures, non-uniformity in compatible species, and differing levels of biological complexity (i.e., populations versus whole communities), no single biotic index can fully quantify the various dimensions of ecological quality needed to inform management decisions. It is thus important – especially when adopting new analytical protocols – to employ multiple index methods simultaneously so that their individual strengths and weaknesses might be evaluated for complementary patterns (Martínez-Crego et al., 2010; Berthelsen et al., 2018; Lu et al., 2021).

The main disadvantage to BRI is its limited geographical scope. While it is widely adopted by municipal and regional monitoring agencies in Southern California (e.g., Gillett et al., 2017; City of San Diego, 2022; LACSD, 2022; OCSD, 2023), adapting it for other regions has proven difficult. Calibrating a pollution gradient from which to derive taxon-specific tolerance values requires a large amount of methodologically-consistent and regionally-complete data on benthic infauna and sediment conditions that few places can offer. Nevertheless, Ranasinghe et al. (2013) calibrated and assessed the performance of BRI, AMBI, and three other index methods for their use in regional monitoring of Puget Sound, where over two decades of data were available by that time. While these indices have not been further validated nor widely applied in Puget Sound, this case study exemplifies how locally- and regionally-developed indices can be adapted for use in other settings.

Along all three of the isobaths assessed in this study, raw bivalve-BRI scores tended to overestimate EQS for samples with relatively low whole-BRI scores (Figure 3.3). This result was anticipated as the mean pollution tolerance values (the variable pi used to calculate BRI; see methods) assigned to bivalve taxa are higher than the combined mean of all other clades for the three species lists (Figure F.1; Appendix F). Bivalve pollution tolerance values were 35.35 for the shallow depth list, 43.27 for the middle depth list, and 42.27 for the deep depth list. Polychaetes (the most numerically abundant and species-rich clade) had mean

values of 32.14, for the shallow depth list, 29.04 for the middle depth list, and 17.98 for the deep depth list. Replicating this study's analyses with polychaetes as the faunal subset resulted in a notably stronger correlation with whole-BRI scores compared to using bivalves (Figure F.4; Appendix F). Malacostracans (the third most abundant clade after bivalves, but the second most species-rich) had mean pollution tolerance values of 14.56, 7.85, and -6.52 for the same respective depth lists. As expected, calculating the malacostracan-BRI scores tended to underestimate whole-BRI scores (Figure F.5; Appendix F). Subset selection should thus be carefully considered when the pollution tolerance values of taxa within that subset skew towards relatively sensitive or tolerant compared to all other infaunal taxa. Subsets with a more uniform distribution of pollution tolerance scores are likely to serve as the best surrogate for the whole fauna in the absence of a regression-based adjustment.

The main disadvantage to AMBI – yet also an attractive feature – is the simplicity of binning taxa into discrete EGs versus assigning unique tolerance values. While the species list for AMBI continues to expand and include more taxa from around the world, it is vital to consider whether the EG assignments therein should be modified to better reflect the nature of a given region or system (Borja and Muxika, 2005). For example, the chemosymbiotic bivalve *Parvilucina tenuisculpta* is assigned to EGIII in the most recent AMBI species list (June 2022), making it considered moderately-tolerant of polluted conditions. However, *P. tenuisculpta* densities were particularly high at the onset of benthic monitoring in the 1970s, and the population crashed after the implementation of partial secondary water treatment in the mid-1980s (Fabrikant, 1984; Swartz et al., 1986; Leonard-Pingel et al., 2019; Kokesh et al., 2022b). The propensity for *P. tenuisculpta* to respond to pollution stress as a first- or second-order opportunist in US waters justifies its reassignment to EGIV on the US-AMBI species list (Gillett et al., 2015). In fact, rerunning the AMBI analyses while referencing the international standard list rather than the US list yielded whole-AMBI and bivalve-AMBI scores that were compressed towards the median of each index's distribution. This is due to

the international standard species list assigning more of the local taxa to moderate ecological groups (EGII and EGIII), while the regionally-adapted US list assigns more taxa to the most sensitive (EGI) and most tolerant groups (EGIV and EGV).

In contrast to bivalve-BRI, raw bivalve-AMBI scores exhibited a tendency to overestimate EQS along each isobath, and adjusting for linear regression had a relatively small impact on agreement (Figures 3.3-3.5). This result is corroborated by the skewed proportion of bivalve taxa assigned to sensitive ecological groups, and lack thereof assigned to tolerant EGs, compared to the whole fauna. 80% of all macrobenthic taxa on the Palos Verdes shelf that are included on the AMBI species list fall into EGI or EGII (Figure F.2; Appendix F). Among the three most abundant classes, these sensitive groups are represented by 71% of polychaete taxa, 89% of bivalves, and 91% of malacostracans. Bivalves (and especially malacostracans) thus largely fail to detect 'Poor' EQS unless the few taxa that fall into EGIV and EGV are well-represented in a sample. The only bivalve species assigned to EGIV and EGV, respectively, are the chemosymbiotic species *P. tenuisculpta* and *Solemya pervernicosa*. However, the binary EQS threshold sufficiently coarsens the requirement for there to be agreement with the whole fauna for >80% of all samples without the need to adjust bivalve-AMBI scores, whereas bivalve-BRI scores only agreed with the whole fauna for 58% of samples prior to adjustment (Figure 3.5). While AMBI is generally regarded to be a poor index in US waters due to an additional apparent bias in performance related to salinity gradients (Gillett et al., 2015) – which may impact the Palos Verdes shelf due to the input of freshwater from the wastewater outfall – it is nevertheless an easily-quantified addition to analytical frameworks under controlled environmental conditions and may serve as an anchor for comparisons with other systems worldwide.

3.5.2 Taxonomic surrogacy: Bivalves are a strong proxy for the living whole-fauna EQS

Using a faunal subset to serve as a proxy for the whole fauna is an example of taxonomic surrogacy (Włodarska-Kowalczuk and Kędra, 2007; Moreno et al., 2007; Bevilacqua et al., 2009; Kokesh et al., 2022b). Because biological monitoring of diverse benthic communities is labor intensive and requires a high degree of taxonomic expertise, prioritizing efforts on a faunal subset can leverage the costs and demands of establishing new programs in poorly-studied regions. Innumerable studies have already investigated the nature of faunal subsets with respect to conventional ecological metrics such as richness, evenness, and beta-diversity (for meta-analyses, see Mellin et al. 2011; Westgate et al., 2014). The reliability of a faunal subset to approximate whole-fauna community patterns appears to be related to the proportional representation of that subset within the whole fauna (Dietl et al., 2016; Kokesh et al., 2022a), making polychaetes and bivalves – or mollusks more broadly – typically the best subsets to employ in most marine settings.

To my knowledge, this study constitutes the first test of robustness for the estimation of BRI scores from a faunal subset (as well, of course, for the counterpart death assemblages). However, faunal subsets have been used in analyses of AMBI applied to polychaetes (Omena et al., 2012) or the indices AMBI, M-AMBI (Muxika et al., 2007), and BENTIX applied to mollusks (Nerlović et al., 2011; Leshno et al., 2016). Importantly, these studies did not compare their subset-generated results to those generated from a more complete whole fauna from the same sampling events. Using data drawn from multiple benthic surveys of European waters, Dietl et al. (2016) found positive correlations between AMBI scores generated using the whole fauna versus using only the mollusk subset. Adjusting mollusk-generated scores for a different set of samples based on the linear regression of the first set resulted in 78% agreement in EQS assignments. The markedly-improved agreement of bivalve-BRI demonstrated here – from 58% for raw scores to 83% for adjusted scores –

similarly demonstrates a strong reflection of EQS by this faunal subset. Additional validation techniques could further indicate various aspects of the fidelity between scores generated by bivalves and the whole fauna, such as recall and precision metrics based on a Bayesian framework (e.g., Pruden et al., 2021).

This study differs from Dietl et al. (2016) in that I used the regression equations to directly adjust the same datasets used to generate those equations, rather than apply those regressions to independent samples via cross-validation. While bootstrapping the linear regression coefficients from which I adjusted bivalve-generated scores partially alleviated this non-independence between the model and the samples assessed (Kohavi, 1995), future work should focus on incorporating monitoring data from elsewhere in the Southern California Bight. Such a compiled dataset would enable the parsing of samples via cross-validation into two analytical groups as performed by Dietl et al. (2016): one from which regression equations can be calculated, and the second that those regressions could be independently applied to for testing agreement of EQS. That way, regression equations that apply throughout the region may be used to facilitate adjusting bivalve-generated scores when the rest of the fauna is not also sampled. Nevertheless, the increased agreement in EQS observed here after applying the adjustments to raw bivalve-BRI scores (Figure 3.5) demonstrates that transformation is warranted for the practical management of benthic ecosystems informed by EQS.

3.5.3 Taphonomic processes: dead bivalves preserve a short memory of past EQS

Much to my surprise, bivalve death assemblages sampled in 2008 and 2016 indicated predominantly 'Good' EQS across the Palos Verdes shelf, with mean BRI and AMBI scores comparable to or better than those of the living community from the 2000s-2010s (Figures 3.6-3.9). Death assemblages thus demonstrated the strongest agreement in ecological qual-

ity with live-collected bivalve samples collected during the same decades: the mean 2008 death assemblage EQS was most similar to that of living bivalves from the 2000s, and the mean 2016 death assemblage EQS was most similar to that of living bivalves from the 2010s. These low index scores for death assemblages are in stark contrast to the expectation that the high-scoring cohorts from peak pollution in the 1970s-1980s would still be part of the death assemblage and thus influenced its score, making it higher. These results of very low live-dead discordance in EQS scores in an area with known strong pollution gradient – and significant live-dead discordance using other metrics – are thus contrary to findings elsewhere that death assemblages typically do include a strong signal of recent past pollution, even using EQS metrics (e.g., Dietl et al., 2016; Leshno et al., 2016; Tweitmann and Dietl, 2018; Aslan and Ovalis, 2023). Here, the availability of a long time-series – five decades of annual data – on the composition of the living bivalve assemblage makes it clear that the death assemblage is failing to detect past pollution: most live-dead studies are based on only a single sampling of living and death assemblages, and it is also rare to have multiple samplings of the co-occurring death assemblages to detect how they might change (Kidwell, 2013).

The age distribution of a death assemblage can be determined by age-dating individual dead shells via geochronological techniques (e.g., radiocarbon or amino acid racemization; Kaufman and Manley, 1998). Like molluscan shell assemblages from warm-temperate and sub-tropical shelf settings elsewhere (e.g., Meldahl et al., 1997; Kidwell et al., 2005; Ritter et al., 2017; Albano et al., 2020), dead bivalve shells (mostly of pollution-tolerant *P. tenuisculpta* and pollution-sensitive, shallow-water-dwelling *Nuculana taphria*) from the Southern California shelf each form a distinct L-shaped, right-skewed distributions where most shells are from recent decades, but a long tail of older shells extend as far back as 10 ka (Tomašových et al., 2014; 2019). The death assemblages sampled in 2008 and 2016 from the Palos Verdes shelf should thus preserve a time-averaged mixture of bivalve shells produced during three important intervals:

1. A '*pre-pollution*' interval from before 1937 when the outfall opened (Figure 3.2). Based on ordination of bivalve assemblages extracted from an independently age-calibrated core (Leonard-Pingel et al., 2019), the pre-pollution bivalve community was likely trophically-diverse (both infaunal and epifaunal suspension feeders, and infaunal deposit feeders, mixed feeders, and chemosymbionts). These assemblages were compositionally similar to living communities observed on this shelf from the 2000s-2010s. Some shells dating to this interval have remained within the upper part of the surface mixed layer and accessible via van Veen grab samplers owing to bioturbation and perhaps being diagenetically-stabilized against further disintegration (Tomašových et al., 2014; 2023).

2. A '*peak-pollution*' interval representing the ∼50 years characterized by rising and initially-declining wastewater emission curve (1940s-1980s; Figure 3.2). Based on biomonitoring, bivalve assemblage compositions were strongly dominated by pollution-tolerant chemosymbiotic taxa (e.g., Lucinidae, Thyasiridae, and Solemyidae), and, as pollution abated and chemosymbionts declined, and increasing proportion of mixed-feeding Tellinidae (Stull et al., 1996).

3. A '*post-pollution*' interval starting when wastewater emissions dropped sharply with the implementation of partial and secondary treatment in the 1980s, with a second strong decline in emissions upon the onset of full secondary treatment in the early 2000s (Figure 3.2). During this interval, the bivalve community became compositionally quite equitable and trophically diverse, more like core assemblages dating to the pre-pollution interval (Figure 4 of Leonard-Pingel et al., 2019). The death assemblages from 2008 and 2016 were sampled during this current, post-pollution interval.

It is difficult to explain why, using EQS metrics, these death assemblages collected late in a 20^{th} Century history of rising and declining pollution do not show scores that are more

intermediate between the 'Good' conditions of today and the 'Poor' conditions of only ∼50 years ago (1960s-1980s). The low BRI and AMBI scores (signaling 'Good' EQS) calculated from death assemblages must reflect a high representation of recently-produced shells from the post pollution interval. Although living bivalves occurred in relatively low densities during the 1990s to 2010s (Table 3.2; also see Leonard-Pingel et al., 2019), these recent cohorts of shell input have had little elapsed time in the surface mixed layer for disintegration, leaving them to dominate the death assemblage. It is possible that the 'Good' EQS from both death assemblages also reflects the inclusion of at least some shells from the compositionally similar pre-pollution interval, since bioturbation admixes shells vertically within the mixed layer and some of these older shells have likely been stabilized diagenetically (per Tomašových et al., 2014; 2023). However, it is curious that the death assemblage would not then also retain a signal from the more recent peak-pollution interval, when huge populations of *P. tenuisculpta* – whose shells are not particularly fragile – would have entered the seabed. (Swartz et al., 1986; Leonard-Pingel et al., 2019; Kokesh et al., 2022b).

Some possible explanations, which are not mutually exclusive, for the apparent lack of peak-pollution shells in modern death assemblages include: (1) the high initial disintegration rate documented for this shelf (Tomašových et al., 2014; 2016) has removed most shells from the relatively-short 40-year interval of peak pollution (unlikely, unless *P. tenuisculpta* has higher-than-average disintegration rates; argued against by shell age-dating, which finds many shells dating to the mid-20^{th} century; Tomašových 2019); (2) an especially high bioturbation or low secondary disintegration rate maintained many shells from the pre-pollution interval up within the surface mixed layer (but it is unclear why bioturbation should favor shells from this interval); (3) preservational bias against the key pollution-tolerant but small-bodied taxa, namely *S. pervernicosa* (likely; very fragile, but quite rare even alive) and *P. tenuisculpta* (unlikely; is not particularly fragile and was wildly abundant alive); and (4) a higher-than-expected production rate of new shells during the post-pollution interval has

overwhelmed the residual peak-pollution signal (argued against by the low densities of living bivalves in the post-pollution interval; Table 3.2). However, despite index scores generated on death assemblages not conforming to expectations, the median proportional abundances of *P. tenuisculpta* in death assemblages were higher than those alive from the 2000s-2010s this metric was strongly correlated to both bivalve-BRI and bivalve-AMBI (Figures F.6-F.7; Appendix F).

There was a small, but qualitatively distinct, improvement in ecological quality (lowering of index scores) observed between the 2008 and 2016 death assemblages (Figures 3.6-3.9). The 2008 death assemblage, forming closer to the peak-pollution interval, should retain a higher proportion of shells produced during or just after the peak-pollution interval; the 2016 death assemblage would reflect an additional eight years of thinning of those cohorts (i.e., loss to disintegration and/or deeper burial) as well as eight years of input from increasingly healthy post-pollution populations. This phenomenon of the composition of a death assemblage lagging behind that of changes in the living assemblage is referred to as 'taphonomic inertia' (Kidwell, 2007; 2008) and is the simplest explanation for why live-dead agreement tends to be lower in settings subject to ongoing or recent anthropogenic disturbance: an age-mixed death assemblage takes time to "catch up" with the new state of the living. These lags are typically estimated to be decadal in scale, as judged from either knowledge of the history of human stressors in the study area or age-dating of the shells of taxa that occur dead-only (e.g., Feser and Miller, 2014; Leshno et al., 2015; Tomašových and Kidwell, 2017; Gilad et al., 2018; Kokesh and Stemann, 2023; Meadows et al., 2023). Conformation of this effect of taphonomic inertia on indices of EQS would be found if dead-shell archives from samples collected along this shelf during the 1970s-1980s exhibited higher BRI/AMBI scores than the death assemblages from 2008.

3.5.4 Additional considerations and confounding factors

Abundances versus biomass: Species richness and numerical abundances are probably the metrics used most commonly to describe spatiotemporal patterns of benthic communities, but the total biomass of sampled organisms has also had a place in benthic ecology for decades: e.g., the species-abundance-biomass (SAB) curves of Pearson and Rosenberg (1978) and the abundance/biomass comparison (ABC) method of Warwick et al. (1987). In practice, disturbed locations are expected to support high-abundances of relatively few small-bodied species, while pristine areas should be capable of supporting more diverse communities that include large-bodied species more sensitive to disturbance. Among the four major macrobenthic phyla that live in soft-bottom habitats, average body weights among species are smallest for annelids (including polychaetes) and progressively higher for mollusks (bivalves), crustaceans (malacostracans), and echinoderms (Warwick and Clark, 1994); the reverse is the case for mean numerical abundances/densities found in sediment grabs. Although polychaetes numerically dominate samples on the Palos Verdes shelf (Kokesh et al., 2022b), the larger-bodied bivalves that place second in abundances account for a disproportionately large portion of the community's total biomass. Substituting biomass for abundances in previous live-dead studies has demonstrated increased sensitivity among samples to variation in primary production, hydrography, and climate change (Powell et al., 1985; Staff et al., 1985; Meadows et al., 2019). Biomass can also readily replace abundances in the calculation of biotic indices, as several studies have demonstrated for AMBI (e.g., Muxika et al., 2012; Cai et al., 2013; Mistri and Munari, 2015). Using biomass data to calculate bivalve-generated BRI and AMBI scores may provide increased sensitivity to spatiotemporal patterns in ecological quality and improve EQS classification accuracy. For example, the documented increase in body size of *P. tenuisculpta* during the peak-pollution interval (Fabrikant, 1984; Stull et al., 1996), which appears to be consistent for specimens found in death assemblages, may increase the influence this species has on bivalve-generated index scores based on biomass,

even though scores calculated with abundances appear to retain little memory from this peak-pollution interval.

Analytical time averaging of living time series: Although live-dead comparisons have recently been used to assess deviation in living communities from the time-averaged condition of an area, they initially served as testing grounds for mechanisms that shape the fossil record and the paleoecological interpretations therein (Behrensmeyer and Kidwell, 1985). However, the difference in the temporal resolution – i.e., the temporal 'pixel size' – between time-averaged death assemblages and snapshot samples of living communities makes the ecological information they possess incongruent: death assemblages can capture higher richness (and by extension, more rare taxa), higher evenness, and lower beta diversity than living communities solely as a consequence of time averaging (Kidwell and Tomašových, 2013). To make live-dead comparisons more temporally-congruent, sequentially-acquired samples of the living should be analytically aggregated into coarser temporal bins, simulating a 'time-summed' assemblage (e.g., Peterson, 1976; Fürsich and Aberhan, 1990; Kowalewski et al., 1998). The exceptionally-long 50-year time series of benthos composition on the Palos Verdes shelf provides an opportunity for assessing the taphonomic processes that would have given rise to the dead-shell compositions sampled in 2008 and 2016. Further, rate parameters for shell disintegration/sequestration/burial derived from age-frequency distributions of bivalve shells along this shelf (Tomašových et al., 2014; 2016; 2023) can be fitted to predict the mechanisms of shell persistence that likely shaped these assemblage compositions.

The Multivariate-AMBI: Developed by Muxika et al. (2007), the multivariate extension of the standard AMBI (M-AMBI) is a popular index that has been increasingly adopted by monitoring programs around the world as a means of classifying EQS (Borja et al., 2019; Pelletier and Charpentier, 2023). In brief, M-AMBI conducts a factor analysis on the normalized AMBI score, species richness, and Shannon diversity of each sample in a dataset and projects these samples onto an orthogonal gradient from user-defined endmember condi-

tions representing fictitious worst-case ('Bad') and best-case ('High') values for each metric. Because death assemblages aggregate the remains of once-living cohorts over many generations, they often capture a higher proportion of an area's species pool than co-occurring living assemblages of the same taxonomic clade (Staff and Powell, 1988; Tomašových and Kidwell, 2009). Thus, others have proposed that death assemblages may be effective baselines from which to define the 'High' condition used to compute M-AMBI (Smith et al., 2023). However, that death assemblages represent the 'pre-impact' reference conditions of a region – especially in settings like this shelf with a long history of human disturbances – should not inherently be assumed. Further, preliminary tests of M-AMBI with the data from this study suggest that separate reference conditions must be set for each of the isobaths to prevent deep-water samples from underestimating EQS (Figure F.8; Appendix F). Although developing protocols for the calculation of M-AMBI for live-dead studies with consideration of taphonomic processes and natural environmental gradients is an important direction for further research, it is beyond the scope of the present study's stated goal of detecting spatiotemporal patterns on a historically-polluted shelf, for which the diversity-independent BRI and AMBI index methods are adequately capable.

3.6 Conclusions

The increasing need to evaluate the ecological quality of coastal benthic habitats, many lacking long-term monitoring histories or perhaps lacking any thorough one-time survey, has made biotic indices an attractive tool for guiding management decisions. The incorporation of dead-shell assemblages or other geohistorical data in these index methods is, in principle, a way to provide reliable estimates of what ecological conditions were like "before now" and, moreover, how conditions have changed by testing for discordance between conditions now and those signaled by the time-averaged death assemblage stressors. Here, I assessed the power of live- and dead-collected bivalves as a surrogate for the entire macroinvertebrate

community on the Palos Verdes shelf in Southern California, a historically acutely-polluted area, and specifically for the biotic indices known as BRI and AMBI.

While scores generated from bivalves alone were positively correlated with scores generated from the whole fauna for both index methods, each index had contrasting strengths. BRI resulted in a tighter correlation between bivalves and the whole fauna, but the bivalve-generated scores consistently overestimated EQS unless they were adjusted by the coefficients of the linear regression; the feasibility of applying these linear regressions more widely in the Southern California Bight requires verification. Correlations for AMBI were weaker, but the concentration of samples towards low values yielded higher agreement in EQS without the need to transform bivalve-generated scores. Due to the well-recognized limitations of biotic index methods in general – each has both strengths and weaknesses – multiple indices should be used simultaneously.

In contrast, applying BRI and AMBI to bivalve death assemblages collected in 2008 and 2016 yielded scores that were most comparable to those of living bivalves sampled from the same decade (2000s and 2010s). Given the time-averaged nature of death assemblages, these results are in contrast to the expectation of higher scores that would have generated if the high production of pollution-tolerant taxa during the 1940s-1980s had been incorporated into the dead. Unlike other studies, here the more recent influx of shells from the two succeeding decades after this interval may be overshadowing these signals, as corroborated by the continuing decline of index scores observed between death assemblages from 2008 to 2016. While death assemblages are an inexpensive and easily-obtainable means of estimating past ecological conditions where time series and other data sources are unavailable, this study exemplifies the critical role of such time series data in verifying the power of death assemblages as surrogates for this history.

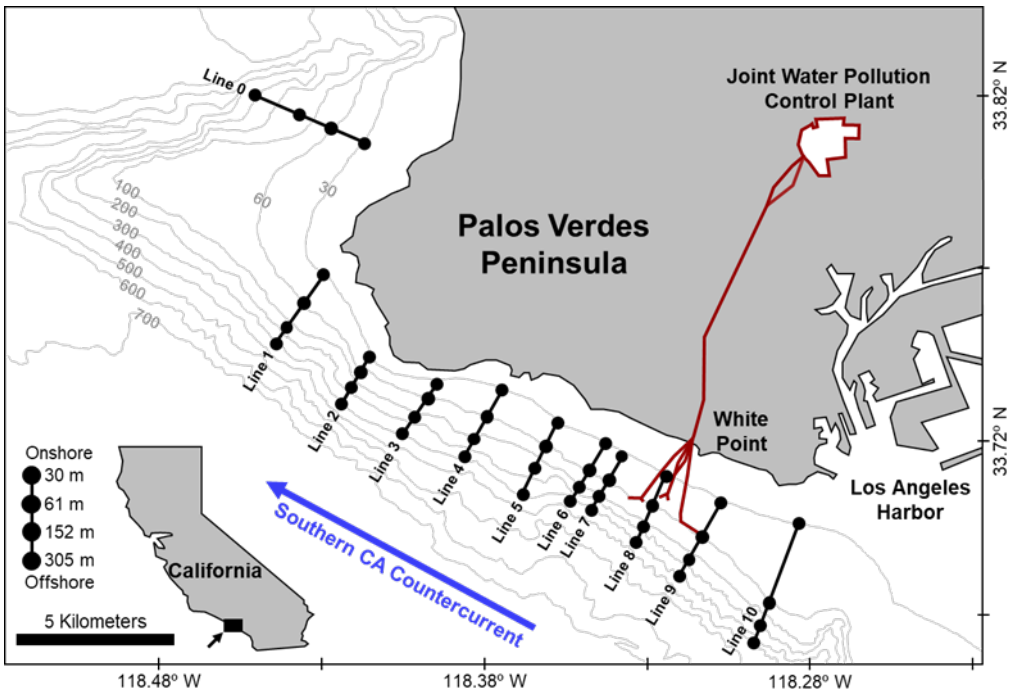

Figure 3.1: Study area off the Palos Verdes peninsula on the southern California continental shelf in Los Angeles County, U.S.A. Macrobenthos are sampled annually at four sites along each of 11 bathymetric transects (Lines 0-10 from north to south) at depths of 30, 61, 152, and 305 m, for a total grid of 44 stations. We exclude stations from the 305 m isobath due to mostly very small dead-shell abundances there. Thin gray lines are isobathic contours in meters and red lines denote the wastewater outfall pipe network extending from the Joint Water Pollution Control Plant (JWPCP) to a series of diffusers in ∼60-m water offshore of White Point. The Southern California Countercurrent flows NW along the shelf (blue arrow). Modified after LACSD (2022).

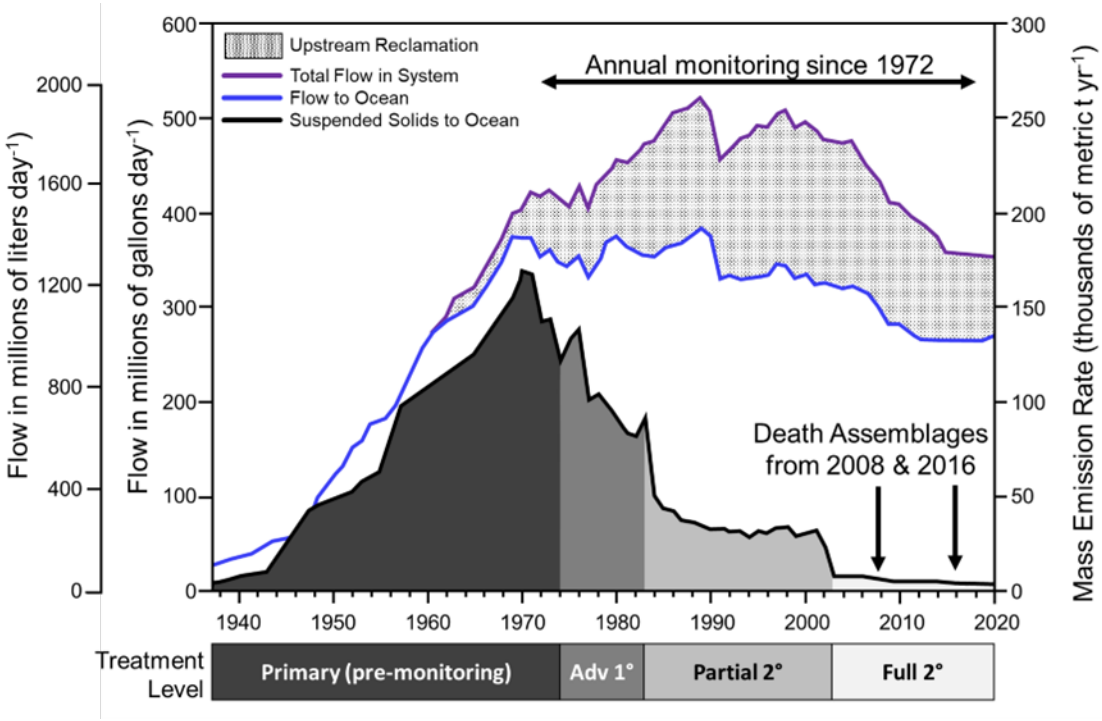

Figure 3.2: History of wastewater emissions to the Palos Verdes shelf from the Joint Water Pollution Control Plant (JWPCP) through the White Point outfall system from 1937-2019. Suspended solid release increased steadily until enactment of the US Clean Water Act in the early 1970s, and then declined with the successive onset of advanced primary wastewater treatment, partial secondary treatment, and full secondary treatment. Annual monitoring of macrobenthic infauna began in 1972, yielding five decades of annually-sampled macrobenthic data. Bivalve death assemblages were collected in 2008. Modified after LACSD (2022).

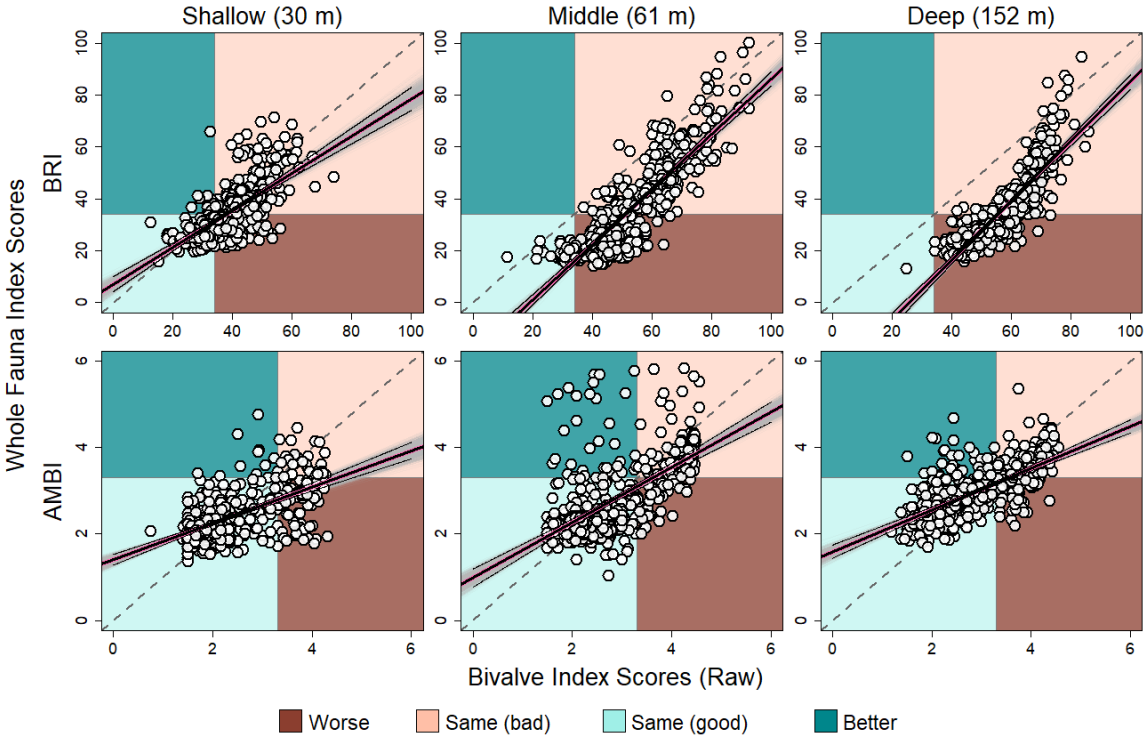

Figure 3.3: BRI (top row) and AMBI (bottom row) scores generated using the whole fauna versus scores generated using only bivalves. Each point represents an annually-collected sediment grab from the 50-year history of benthic sampling from 11 stations along three isobaths (shallow to deep; left to right) on the Palos Verdes shelf. The dashed grey line identifies the 1:1 relationship. Shaded backgrounds indicate the type of agreement in EQS of points plotted within those areas: the blue area in the lower-left quadrant represents agreement of 'Good' status (i.e., requires no remediation), the pink area in the upper-right quadrant represents agreement of 'Poor' status (i.e., requires remediation), the dark teal area in the upper-left quadrant represents samples for which bivalves indicate 'Worse' status (i.e., bivalves indicate 'Poor', but the whole fauna indicates 'Good'), and the brown area in the lower-right quadrant represents samples for which bivalves indicate 'Better' status (i.e., bivalves indicate 'Good', but the whole fauna indicates 'Poor'). The solid black lines represent the trendline generated by the mean coefficients of these bootstrapped trendlines and their 95% confidence intervals, and the pink line is the linear trendline of the full dataset. For BRI, regression intercepts became increasingly negative from the shallow to deep datasets, resulting in bivalve-BRI scores that accurately estimate the status of whole-BRI in areas with poor health (points in the pink 'Poor' status agreement quadrant), but tending to underestimate the quality of benthic health under healthy conditions (the brown 'Worse' status quadrant). In contrast, the relatively high intercepts for AMBI regressions resulted in many bivalve-AMBI scores sufficiently estimating whole-AMBI status in healthy areas (i.e., points in the 'Good' status agreement quadrant), but they tended to overestimate EQS under poor conditions (the teal 'Better' status quadrant).

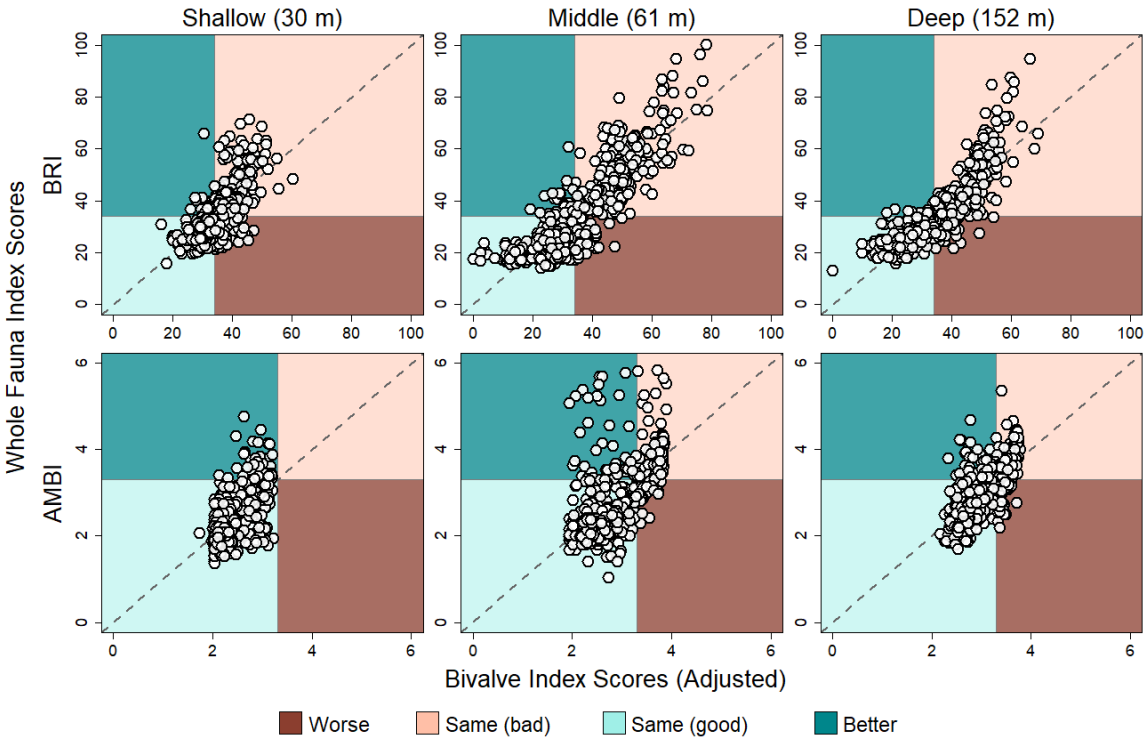

Figure 3.4: The relationship between BRI (top row) and AMBI (bottom row) scores generated using the whole fauna versus scores generated using only bivalves, as presented in Figure 3.3. Bivalve-generated scores were adjusted according to the mean bootstrapped linear coefficients calculated from the directly-calculated raw scores presented in Table 3.3. Note that scores generated by the whole fauna along the y-axis have not been altered. For BRI, adjusting bivalve-generated scores shifted points towards the left of the x-axes, enabling more points to occupy regions of status agreement (blue and pink areas; lower-left and upper-right quadrants). In contrast, adjustment of bivalve-AMBI scores caused points to compress along the x-axis, resulting in bivalves largely failing to detect 'Poor' EQS (pink and brown areas; right-side of plots).

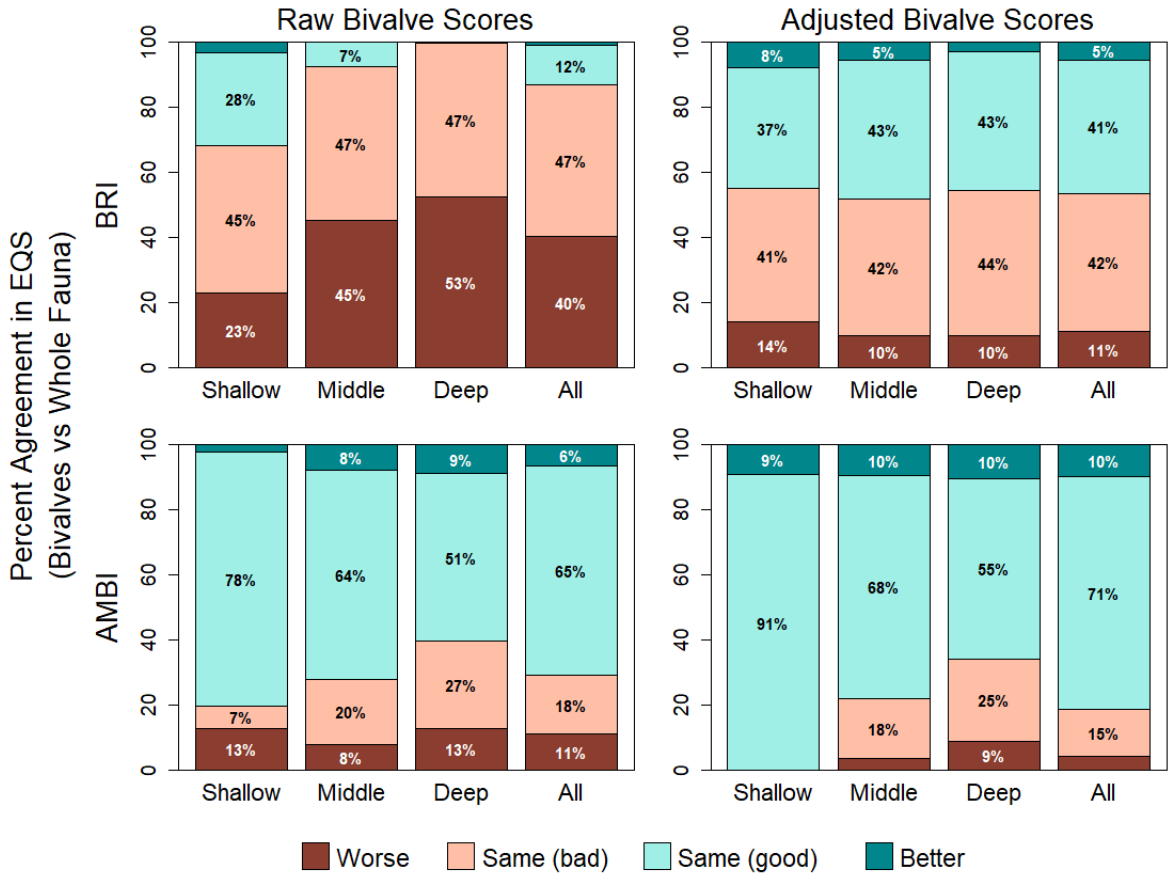

Figure 3.5: Proportional agreement of EQS assignments from BRI (top row) and AMBI scores (bottom row) generated from the whole fauna and either directly-calculated raw bivalve scores (left column) and after adjusting bivalve scores based on bootstrapped linear regressions (right column). Bar colors indicate whether bivalve-generated scores resulted in (1) a 'Worse' EQS than the whole fauna (brown segments, i.e., bivalve status = 'Poor' while whole fauna status = 'Good'), (2) a 'Better' EQS (dark teal; i.e., bivalve status = 'Good' while whole fauna status = 'Poor'), or (3) the same EQS as the whole fauna (pink bars when both = 'Poor', blue bars when both = 'Good'). Percentage values are printed for bar segments that round to at least 5%. The agreement between raw bivalve-BRI and whole-BRI scores decreased from the shallow to deep isobaths, but agreement at all isobaths notably improved after scores were adjusted. In contrast, raw bivalve-AMBI demonstrated relatively high agreement with whole-AMBI at all depths, although almost exclusively for samples where EQS was 'Poor'. Bivalve-AMBI agreement with whole-AMBI remained largely unchanged after scores were adjusted, but adjusted bivalve-AMBI scores failed to detect any samples as exhibiting 'Poor' EQS along the shallow isobath.

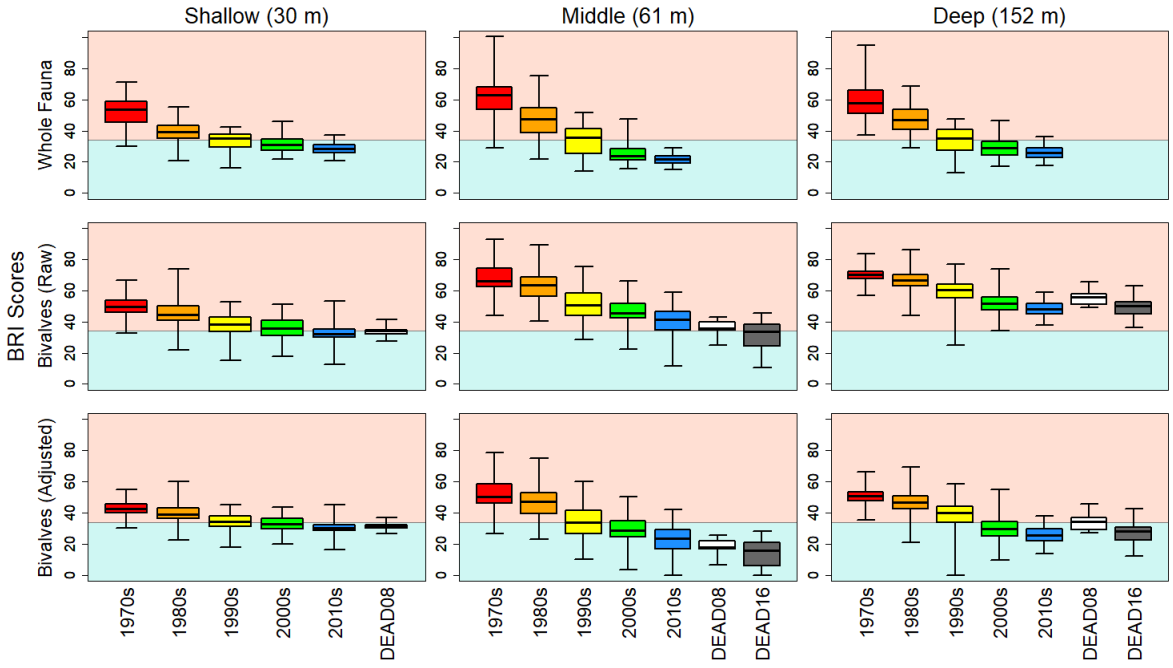

Figure 3.6: Boxplots of whole-BRI (top row), raw bivalve-BRI (middle row), and adjusted bivalve-BRI scores based on bootstrapped linear regressions (bottom row). White and dark boxes represent bivalve death assemblages sampled in 2008 and 2016, respectively. BRI scores declined steadily over time, transitioning from predominantly 'Poor' (pink shaded area) to 'Good' EQS (blue shaded area). Bivalve death assemblages yielded predominantly 'Good' adjusted EQS (similar to living bivalve scores from the 2000s and 2010s), and BRI scores were even lower for 2016 death assemblages compared to those from 2008.

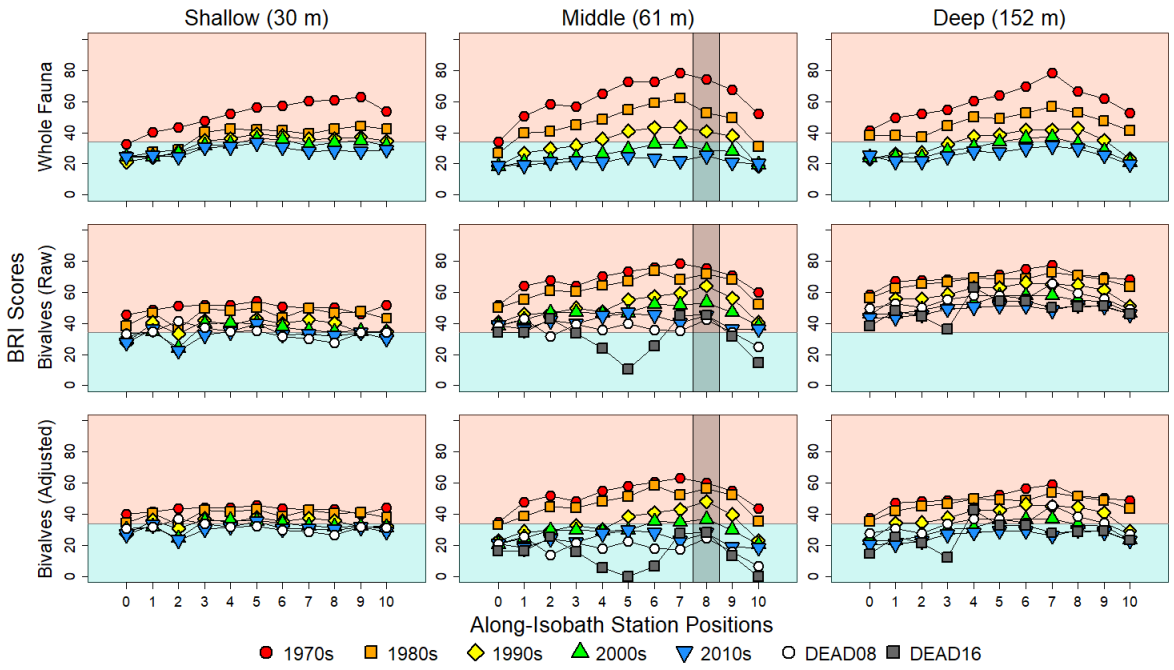

Figure 3.7: Spatiotemporal variation in BRI scores along the Palos Verdes shelf, with columns for water depth and rows for taxonomic sets organized as in Figure 3.6. The x-axis of each plot represents sampling stations along their respective isobaths (onshore-offshore transect Line 0 from the north to Line 10 to the south; Figure 3.1). The outfall openings off of White Point are approximately located at 60 m depth near Line 8 and indicated by the grey shaded region in the middle column. White circles and dark squares in the second and third rows are scores generated from bivalve death assemblages sampled in 2008 and 2016, respectively. The highest whole-BRI scores (i.e., poorest ecological quality) are temporally from the first two decades and spatially near Lines 7-8, with conditions improving distally from the outfall source. Whole-BRI scores consistently declined over the decades, and the spatial variation along the shelf was damped by the 2000s (but a faint signal persisted into the 2010s at the deep isobath). Raw bivalve-BRI primarily indicated 'Poor' EQS, but preserved a similar spatiotemporal distribution as whole-BRI. Adjusted bivalve-BRI improved the EQS agreement with whole-BRI such that mean scores from the 2000s-2010s indicated 'Good' EQS. Bivalve death assemblages yielded relatively low and spatially-indistinct scores barring two cases: (1) the 2008 death assemblage from the deep isobath yielded relatively high scores with a similar spatial gradient as the 1970s-1980s living bivalves outward from Line 8, and (2) the 2016 death assemblage from the middle isobath yielded particularly low scores with a strong spatial gradient that peaks at Lines 2 and 8 and dips at Lines 5 and 10. Overall, the relatively low BRI scores generated by death assembles are not consistent with time-averaging over the entire 20^{th} and 21^{st} centuries so far.

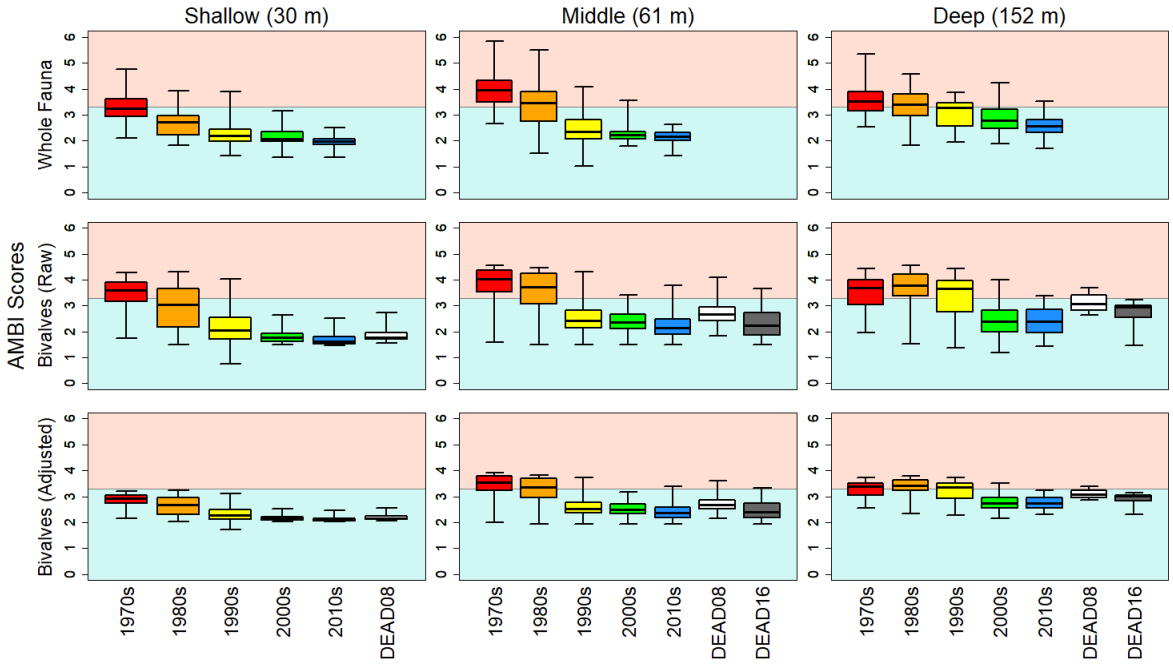

Figure 3.8: Boxplots of whole-AMBI (top row), raw bivalve-AMBI (middle row), and adjusted bivalve-AMBI scores based on bootstrapped linear regressions (bottom row). White and dark boxes represent bivalve death assemblages sampled in 2008 and 2016, respectively. AMBI scores declined steadily over time, transitioning from predominantly 'Poor' (pink shaded area) to 'Good' EQS (blue shaded area). Bivalve death assemblages yielded predominantly 'Good' adjusted EQS (similar to living bivalve scores from the 2000s and 2010s), and AMBI scores were detectably lower for the 2016 death assemblage compared to that from 2008.

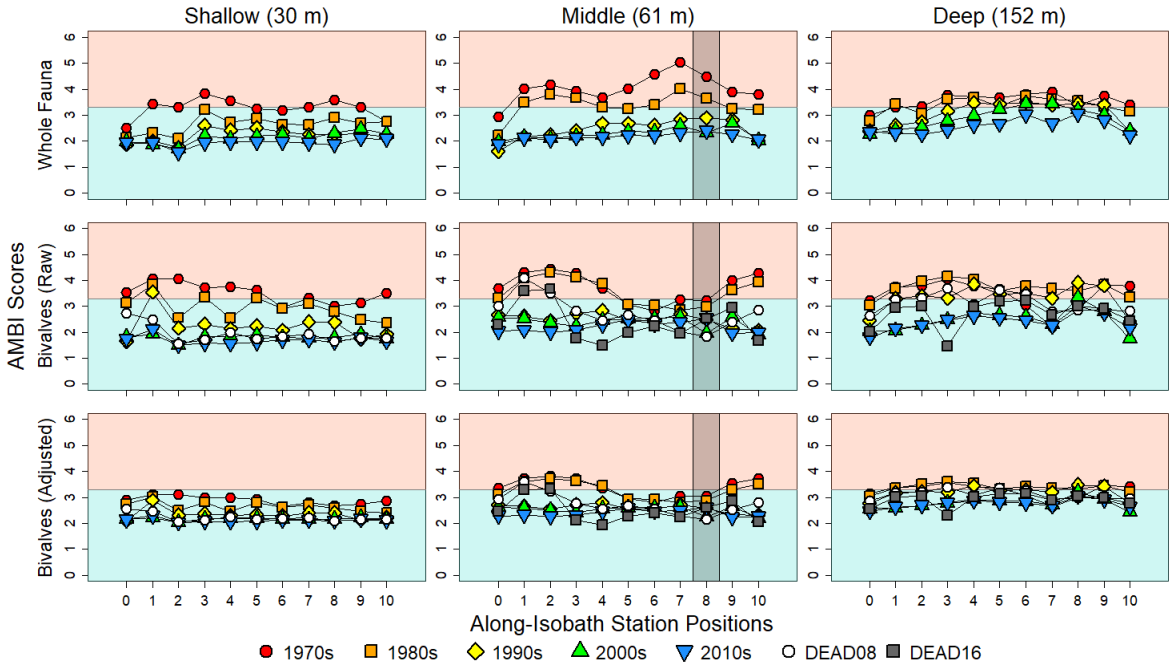

Figure 3.9: Spatiotemporal variation in AMBI scores along the Palos Verdes shelf, with columns for water depth and rows for taxonomic sets organized as in Figure 3.8. The x-axis of each plot represents sampling stations along their respective isobaths (onshore-offshore transect Line 0 from the north to Line 10 to the south; Figure 3.1). The outfall openings off of White Point are approximately located at 60 m depth near Line 8 and indicated by the grey shaded region in the middle column. White circles and dark squares in the second and third rows are scores generated from bivalve death assemblages sampled in 2008 and 2016, respectively. The highest whole-AMBI scores (i.e., poorest ecological quality) are temporally from the first two decades and spatially create two peaks: one near Lines 1-2 and another at Lines 7-8, with conditions improving distally from these locations. Whole-AMBI scores consistently declined over the decades, and the spatial variation along the shelf was damped by the 2000s. Raw bivalve-AMBI primarily indicated 'Good' EQS after the 1980s, and similarly to whole-BRI, peaked near Line 2. Adjusted bivalve-AMBI shifted scores downward, resulting in all mean scores from the shallow isobath falling beneath the threshold value and indicating 'Good' EQS. Bivalve death assemblages yielded moderate scores and notably peaked near Line 2 along the middle isobath. Overall, the relatively moderate AMBI scores generated by death assembles are more consistent than BRI with respect to time-averaging over the entire 20^{th} and 21^{st} centuries so far.

Table 3.1: Ecological quality status (EQS) thresholds and their associated score ranges as defined in this study for the Benthic Response Index (BRI; Smith et al., 2001; 2003) and ATZI's Marine Benthic Index (AMBI; Borja et al., 2000).

	Minimum Score	Maximum Score	EQS	Binary EQS
BRI	No lower bound	25	Reference	Good
	25	34	Marginal Deviation	Good
	34	44	Biodiversity Loss	Poor
	44	72	Community Function Loss	Poor
	72	No upper bound	Defaunation	Poor
AMBI	0.0	1.2	High	Good
	1.2	3.3	Good	Good
	3.3	4.3	Moderate	Poor
	4.3	5.5	Poor	Poor
	5.5	6.0	Bad	Poor

Table 3.2: Mean density of individuals and species richness (in parentheses) per van Veen grab per decade for three water depths on the Palos Verdes shelf (11 stations per isobath sampled annually), calculated for the whole macrobenthic fauna (top set) and bivalves only (middle set); mean density and richness of bivalve death assemblages sampled in 2008 and 2016 given in the two far-right rows (asterisks: note the dead from 2016 are only available for the middle and deep isobaths and not completely sampled). All mean values are rounded to the nearest whole number. Bottom set of numbers are the proportional density and richness of bivalves within the whole fauna, rounded to the nearest tenth of a percent.

	Depth	1970s	1980s	1990s	2000s	2010s	Dead08	Dead16
Whole Fauna	Shallow 30 m	702 (55)	597 (77)	587 104)	673 (127)	664 (122)		
	Middle 61 m	1,122 (39)	820 (57)	586 (83)	608 (103)	435 (102)		
	Deep 152 m	608 (27)	919 (41)	638 (56)	436 (64)	295 (62)		
Bivalves	Shallow 30 m	99 (8)	81 (9)	32 (9)	34 (10)	35 (9)	598 (29)	
	Middle 61 m	402 (6)	319 (8)	50 (9)	35 (9)	18 (8)	365 (20)	*81 (7)
	Deep 152 m	415 (6)	439 (6)	287 (6)	21 (6)	35 (7)	257 (14)	*149 (8)
% Biv. of Whole	Shallow 30 m	14.1 (14.5)	13.6 (11.7)	5.5 (8.7)	5.1 (7.9)	5.3 (7.4)		
	Middle 61 m	35.8 (15.4)	38.9 (14.0)	8.5 (10.8)	5.8 (8.7)	4.1 (7.8)		
	Deep 152 m	68.3 (22.2)	47.8 (14.6)	45 (10.7)	4.8 (9.4)	11.9 (11.3)		

Table 3.3: Linear regression models for BRI (top set) and AMBI scores (bottom set) generated using the whole fauna versus using only bivalves. Linear coefficients (±95% confidence intervals) were calculated for each dataset's average bootstrapped model (i.e., the result of 5,000 resampling-with-replacement events from each isobath's full dataset) and full model (i.e., true regression using all 55 datapoints per isobath). Pearson's correlation coefficients for the full model (all significant; p-values <0.001) are in the far-right column.

		Bootstrapped Model		Full Model		Pearson's r
	Isobath	Slope	Intercept	Slope	Intercept	(all p <0.001)
BRI	Shallow 30 m	0.714 ± 0.080	7.154 ± 3.043	0.714 ± 0.072	7.183 ± 2.920	0.669
	Middle 61 m	1.065 ± 0.068	-20.319 ± 3.494	1.065 ± 0.056	-20.269 ± 3.072	0.856
	Deep 152 m	1.144 ± 0.082	-29.450 ± 4.498	1.145 ± 0.066	-29.511 ± 3.879	0.847
AMBI	Shallow 30 m	0.420 ± 0.059	1.407 ± 0.125	0.421 ± 0.050	1.406 ± 0.124	0.607
	Middle 61 m	0.640 ± 0.066	0.984 ± 0.207	0.640 ± 0.071	0.983 ± 0.209	0.618
	Deep 152 m	0.482 ± 0.048	1.598 ± 0.155	0.483 ± 0.049	1.595 ± 1.595	0.671

CHAPTER 4

BIVALVE DEATH ASSEMBLAGES RECORD DYNAMICS AND CONSEQUENCES OF RECENT BIOLOGICAL INVASIONS IN KINGSTON HARBOUR, JAMAICA

4.1 Abstract

Short-lived biological invasions may leave lasting impacts on ecosystems well after they have concluded, yet the nature of such events is difficult to elucidate in the absence of monitoring efforts. Here, the ability for surficial death assemblages to recount such invasion events and their ecological legacies was tested using mangrove-dwelling bivalves from Kingston Harbour, Jamaica, where the Asian green mussel (*Perna viridis*) was introduced ∼20 years ago. While rare in Kingston Harbour today, relative densities of dead *P. viridis* shells mapped well to historic surveys from early into the invasion and thus help reconstruct spatial variations in invasion intensity. Live-dead discordance of the epifaunal bivalve community further indicated that species have not returned to pre-invasion relative abundance distributions: the economically-important mangrove oyster (*Crassostrea rhizophorae*) has notably declined while the flat tree oyster (*Isognomon alatus*) rose to dominance. Finally, we report the presence of the charru mussel (*Mytella strigata*) in Kingston Harbour, a newly-introduced species that has not yet been significantly incorporated into the subfossil record. This case study exemplifies the utility of underexploited sources of geohistorical data for informing the growing problem of human-assisted biological invasion.

This chapter was originally published as: Kokesh B.S. and Stemann T.A. Dead men still tell tales: bivalve death assemblages record dynamics and consequences of recent biological invasions in Kingston Harbour, Jamaica. *Geological Society, London, Special Publications*, 529(1):65–78, 2023. doi:10.1144/SP529-2022-28.

4.2 Introduction

Human-mediated range expansion of invasive species is an increasingly common, but unintended, consequence of international transportation and trade (Hulme, 2009; Tabak et al., 2017). This is particularly the case for aquatic invertebrates capable of dispersal via ballast water from large vessels, with notable examples including the Eurasian zebra mussel (*Dreissena polymorpha*) introduced to freshwater systems of North America (Strayer, 2009), sea walnut ctenophore (*Mnemiopsis leidyi*) introduced to Europe and West Asia (Shiganova et al., 2019), and the European green crab (*Cancer maenas*) introduced to nearly every continent (Klassen and Locke, 2007). Such invasions have devastating potential to permanently strain and/or restructure biological communities and the frequent infeasibility of eradication requires managers to prioritize containment, suppression, or other alternative control strategies (Green and Grosholz, 2020).

The utility of death assemblages – multi-generational aggregations of skeletal remains within sediments – to evaluate ecological responses to invasive species is well recognized, yet infrequently applied (Dietl and Flessa, 2011; Kidwell and Tomašových, 2013). Conceptually, a recent invader will appear either live-only or underrepresented in the dead, as the death assemblage has not had sufficient time to reflect the composition of the new community ('taphonomic inertia' *sensu* Kidwell, 2008; see also Albano et al., 2021). In contrast, appreciable representation in death assemblages has also been used to indicate that previously recognized invasive species are, in fact, returning natives to the community (Betancourt et al., 1984; Flenley et al., 1991; Coffey et al., 2011). The use of death assemblages to reconstruct the unobserved history of a biological invasion and/or quantify its consequences on communities has been particularly well-studied for invasive gastropods (Yanes, 2012; Chiba and Sato, 2013; Smith and Dietl, 2016; Kusnerik et al., 2020) and more recently with bivalves (Albano et al., 2018).

Here, we test whether the introduction of the Asian green mussel (*Perna viridis*) to

Kingston Harbour, Jamaica ∼20-years ago has left (1) a record of invasion intensity in death assemblages, and (2) lasting consequences on current community composition. *Perna viridis* is native to the coastal waters of the Indo-Pacific, following the coastline from northern China, Indonesia, India, and westward as Iran (Siddall, 1980; Dias et al., 2018) and most frequently occurs in the shallow-subtidal to lower-intertidal zones (Baker et al., 2007). The first west Atlantic sighting of *P. viridis* was reported in 1990 from the port of Point Lisas in Trinidad (Agard et al., 1992). It since spread through the Gulf of Paria across Trinidad and NE Venezuela by 1993 (Rylander et al., 1996), jumped to Kingston Harbour, Jamaica, in 1998 (Buddo et al., 2003), Tampa Bay, Florida, in 1999 (Ingrao et al., 2001; Power et al., 2004), Cienfuegos Bay, Cuba, in 2005 (Fernández-Garcés and Rolán, 2005), and likely elsewhere. The distribution of *P. viridis* across the west Atlantic is considered patchy and limited genetic variation among populations even decades after initial discovery suggests these populations were established after a single introduction that subsequently spread with the aid of human shipping traffic (Baker et al., 2007; Gobin et al., 2013; Gilg et al., 2014). However, the geographic extent of this invasion continues to grow as *P. viridis* has more recently been reported along the coast of Colombia (Gracia and Rangel-Buitrago, 2020) and as far south as Rio de Janeiro, Brazil (de Messano et al., 2019).

Three aspects of the invasion of Jamaica's Kingston Harbour make this event unique: (1) *P. viridis* populations rapidly overtook subtidal surfaces at varying intensities across the harbour (mean densities of 32.46 to 1,427.21 individuals/m^2; Buddo et al., 2003), (2) high growth and colonization rates determined via cage experiments demonstrated that *P. viridis* thrived in the harbour's degraded conditions (Buddo, 2008), and (3) populations significantly declined across Kingston Harbour by 2010. Although not extirpated, the reasons for this 'boom-bust' dynamic are unclear. Further, the direct and indirect ecological impacts of *P. viridis* on the native fauna, plus the nature of recovery by native species, are unexplored. Our methodological approach of comparing mangrove- and wharf-dwelling bivalve communities

to underlying death assemblages composed of allochthonous shell inputs (as opposed to *in-situ* benthic communities and their dead shells) was designed to determine whether death assemblages are a useful compliment to biological surveys that otherwise may not involve sampling from the sediment.

4.3 Methods

4.3.1 Study area

Kingston Harbour is a natural embayment on the southeast coast of Jamaica, covering an area of ∼50 km^2 (Figure 4.1). The harbor is surrounded by the capital city of Kingston to the north, municipalities of Portmore and Harbour View to the west and east, respectively, and the Palisadoes – a tombolo derived from terrigenous sediments connecting a series of keys to the mainland – to the south (Hendry, 1978). Besides providing overland access to the town of Port Royal and other important infrastructure (e.g., Norman Manley International Airport, Jamaica Maritime Institute, etc.), the inner margin of the Palisadoes supports dense mangrove swamps, which in turn, stabilize the tombolo from erosion. Despite decades of physical destruction (Goodbody, 2003), wastewater eutrophication (Webber et al. 2003; Francis et al., 2014), pollution (Rose and Webber, 2019), and introduced species (Buddo et al., 2003), mangrove swamps remain critical habitats for the intertidal and subtidal marine communities of Kingston Harbour, particularly the diverse epibiont biota consisting of molluscs, bryozoans, ascidians, sponges, and crustaceans that live on mangrove prop roots (Alleng, 1997; Elliott et al., 2012).

Samples were collected in early 2019 from 35 stations across five sites in and near Kingston Harbour (Table 4.1). From west to east, Great Salt Pond (GSP; 3 stations) is a semi-enclosed lagoon SW of the entrance to Kingston Harbour. Old Coal Wharf (OCW; 5 stations) is a shallow clearing near Port Royal consisting of a series of wharf pilings. In late 2019,

construction began at OCW for the Port Royal Cruise Port, which may have removed the original pilings from which we sampled. Goodbody Channel (GC; 8 stations) is a small inlet near the mouth of the larger Fort Rocky Lagoon. Refuge Cay (RC; 10 stations) is a large key located north of and disconnected from the Palisadoes. Finally, Buccaneer Swamp (BS; 9 stations) refers to the mangroves and wharfs surrounding the airport, Jamaica Maritime Institute, Royal Jamaica Yacht Club, and Buccaneer Beach. These five sites were targeted as they coincide with stations where previous researchers recorded *P. viridis* densities shortly after their introduction. Criteria for establishing individual stations at each site were (1) at least 5 m of separation, (2) colonization surfaces that extend below the intertidal zone, and (3) above or adjacent to a sufficient volume of sediment for the collection of death assemblages (see below). GSP and OCW had relatively few stations compared to the other sites due to limited availability of biofouling surfaces that reached below the intertidal zone.

4.3.2 Sample collection and processing

At each station, living communities on mangrove prop roots and wharf pilings were surveyed by directly counting the number of all living individuals visible to the naked eye in the field ($\geq$5 mm) among different species. To directly compare results to Buddo et al. (2003), all prop roots within a 1 m^2 area were examined per sample, and a similar area was surveyed on wharf pilings. Death assemblages were collected by exhuming 15,625 cm^3 of sediment (a cube with 25 cm dimensions) directly beneath surveyed mangrove roots or adjacent to the studied face of wharf pilings. Shells were separated from fine sediment in the field using a box sieve with 5 mm mesh, dried on tables in the lab, and counted by species. Although this mesh size avoids the collection of small-bodied and juvenile specimens (Kidwell, 2002), focusing on large individuals was advantageous because (1) we almost certainly failed to count most living small-bodied individuals without resorting to destructive sampling (e.g., scraping communities or breaking branches to bring back to the lab), (2) large specimens

would more likely fall directly into sediments beneath where they lived, minimizing post-mortem transport (although water energy is significantly reduced along the inner margin of the Palisadoes compared to the outer harbour; Sherwin and Deeming, 1980), and (3) large specimens suggest populations were present at a locality for an appreciable window of time to support mature individuals. Although live-collected infaunal and epibenthic bivalve species were also documented, no living individuals of mangrove-dwelling species were found within sediment samples. Thus, we assumed that all dead shells from mangrove-dwelling species represent allochthonous inputs from the overlying prop roots and wharf surfaces.

The final dataset included five mangrove- and wharf-dwelling species from the living and death assemblages (Figure 4.2): *Brachidontes exustus* (native; scorched mussel), *Crassostrea rhizophorae* (native; mangrove oyster), *Isognomon alatus* (native; flat tree oyster), *Mytella strigata* (non-indigenous, previously unreported; *charru mussel*), and *Perna viridis* (non-indigenous, introduced in 1998; green mussel). These taxa comprised a total of 31,994 sampled individuals; 3,309 living and 28,685 dead (Table 4.2). For dead shells, loose disarticulated valves represented by a mostly-complete hinge and paired valves still connected by a ligament were both counted as individuals. Prior to running data analysis, the number of dead specimens was divided by two to correct for potential collection of disarticulated valves of the same individual (Kowalewski and Hoffmeister 2003). Other epibiont bivalves were present in both the living community and death assemblage – namely *Geukensia granosissima*, *Modiolus squamosus*, *Mytilopsis sallei*, *Pinctada imbricata*, and *Plicatula gibbosa* – but none of these taxa exceeded a combined living and dead sample size of 20 individuals. Thus, they were removed from the dataset prior to further analyses (counts are still included in the data file; Appendix).

4.3.3 Data analysis

Mean monthly densities (individuals per m^2) of living *P. viridis* from February 2000 to January 2001 were taken from Buddo et al. (2003) for BS (Old Runway; 'OR' in their study), GSP, and OCW. Estimated densities at GC and RC were taken as the average 122-day colonization density for an experimental station at each site from Buddo (2008). Mean densities for non-zero samples from death assemblages and the 2019 living community were quantified based on abundances per sample from our data. Relative densities across all three datasets ('historic community' from Buddo et al., 2003, dead, and living) were compared using bar plots with Spearman's rank-order correlation tests.

Raw counts of all five bivalve species were converted to proportional abundances and compared between living communities and death assemblages at the resolutions of (1) 35 individual samples, (2) aggregated among 5 sites, and (3) summed for all of Kingston Harbour. The use of proportional abundances was necessary to facilitate comparisons among sites with disparate numbers of samples, as well as between the living community and death assemblage due to different methods and units of sampling. Living and dead proportional abundances were plotted against each other for each species to determine whether species were equally represented (1:1 ratio), overrepresented in the death assemblage (dead proportional abundance $>$ living), or overrepresented in the living community. Live-dead discordance was quantified using Spearman's ranked-order correlation. Differences in living and dead proportional abundance for each species when summed across all of Kingston Harbour were statistically evaluated via bootstrapped resampling (1,000 iterations). The bootstrapping method randomly selects individual samples with replacement 35 times (the number of actual samples), sums the counts for each species, then recomputes proportional abundances. These 1,000 simulated proportional abundances were used to generate 95% confidence intervals and test for significant differences between the living and dead abundances for each species.

Compositional variation among samples was evaluated using non-metric multidimensional scaling (NMDS; Kruskal, 1964) to ordinate samples in low-dimensional space. NMDS ranks pairwise distances over a predefined number of axes (here, two) and heuristically searches for axis scores that minimizes a loss function. We calculated abundance-based Bray-Curtis dissimilarities on Hellinger-transformed (square-root proportional) abundances as input for ordination. Compositional differences between the living and dead for each site was assessed using permutational multivariate analysis of variance (PERMANOVA; Anderson, 2001).

We calculated the multivariate dispersion of living communities and death assemblages among all sites to test whether the degree of compositional variation differed between living and dead. Multivariate dispersion measures the distance between individual samples to their corresponding group centroids in multivariate ordination space and is often treated as a proxy for -diversity (Anderson et al., 2006). We tested whether living communities and death assemblages from each site differed in dispersion but did not test for differences among the living ('live-live' comparisons) and dead ('dead-dead' comparisons) across sites due to disparate sample sizes (e.g., 3 for GSP, 10 for RC). Live-dead differences were tested using a two-sample randomization test (1,000 iterations), which compares the true difference in mean dispersions within sites to that generated when the assemblage labels (live and dead) are randomly reassigned.

All analyses were conducted using R version 4.1.0 (R Core Team, 2023).

4.4 Results

4.4.1 Green mussel densities through time

Densities of *P. viridis* shells reported in the three datasets studied here (historical living community sampled 20 years ago, current living community, and current death assemblage) demonstrated nearly perfect ranked-order correlation among the five studied sites of Kingston

Harbour (Spearman $\rho = +1$ for historic-live, $\rho = 0.9$ for historic-dead or live-dead; Figure 4.3). GSP had the lowest densities in the historical community (32.5 individuals/m^2) and no *P. viridis* were found in either the living community or death assemblage. For the historical community, *P. viridis* densities increased among sites from OCW (279.9 ind/m^2), GC (344.5 ind/m^2), RC (1,193.3 ind/m^2), and BS (1,342.1 ind/m^2). Densities for the current living community dropped by orders of magnitude but remained in perfect ranked order with the historical community starting with OCW (0.2 ind/m^2), GC (0.38 in/m^2), RC (0.4 ind/m^2), and BS (0.8 in/m^2). Death assemblage densities, despite being based on different sampling techniques, were ordered from OCW (3.6 ind/0.015 m^3), GC (3.9 ind/0.015 m^3), BS (14.2 ind/0.015 m^3), and RC (16.3 ind/0.015 m^3). Interestingly, the historical community and death assemblage patterns both feature a large jump in density between GC and RC/BS while the left-to-right increase for the current living community is more stepwise.

4.4.2 Live-dead discordance of proportional abundances

Compositional fidelity between the overall living community and death assemblage demonstrated unique trends for each taxon (Figure 4.4). *Brachidontes exustus*, the second-most abundant taxon in both assemblages, is the only one with similar values based on the percent change from dead to living (a +0.3% increase from dead to living). *Crassostrea rhizophorae* was the dominant bivalve found in the death assemblage, accounting for 57.3% of individuals. However, *C. rhizophorae* demonstrated a -73.8% decrease in the living community, shifting its position to the third-ranked taxon accounting for 15% of individuals. *Isognomon alatus* demonstrated the inverse of *C. rhizophorae*: while ranked third for the death assemblage (13% of individuals), its living community proportional abundance accounted for a +280.6% increase, establishing *I. alatus* as the current dominant species in the community at 49.3% of individuals. *Mytella strigata* and *P. viridis*, the fourth- and fifth-ranked taxa in both assemblages (due to a short window of time-averaging constrained by the time of introduc-

tion), respectively, demonstrated opposing trends: *M. strigata* was overrepresented in the living community compared to the death assemblage (+226.5% change) while *P. viridis* was underrepresented in the living (-81% change). Overall rank-abundance correlation was not significant (Spearman $\rho = 0.6$, p = 0.35), but ρ remained positive as the only trade in ranks occurred between *C. rhizophorae* and *I. alatus*. Consequently, rank-abundance correlation based on only the three native species (i.e., those with long time-averaging windows unconstrained by a time of introduction) was perfectly negative such that ρ = -1. Differences in proportional abundance between the death assemblage and living community for all taxa except *B. exustus* were supported by bootstrapped 95% confidence intervals (not shown in Figure 4.4 due to small size of error bars).

Patterns from the summed proportional abundances for the living communities and death assemblages above are further supported at the resolution of individual samples and sites (Figure 4.5). In particular, the change in community dominance from *C. rhizophora*e to *I. alatus* is apparent in nearly all samples that include both species. Although their proportional abundances are small, the overrepresentation of *P. viridis* and underrepresentation of *M. strigata* in death assemblages were both apparent. Only *B. exustus* and *M. strigata* were recorded at GSP (diamonds; Figure 4.5), the former dominating the death assemblage, and the latter dominating the living community.

4.4.3 Variation in composition and dispersion

Ordination revealed spatial and temporal variation in faunal composition (Figure 4.6). Spatially, samples from GSP were isolated from the rest of the samples in the negative direction along NMDS1, likely defined by the occupation of only *B. exustus* and *M. strigata* at GSP. In contrast, samples from the four inner harbour sites (BS, GC, OCW, and RC) all overlapped on the positive end of NMDS1 where *C. rhizophorae*, *I. alatus*, and *P. viridis* were present, but *M. strigata* was less abundant. Temporally, living communities and death assemblages

formed adjacent (but minimally overlapping) clusters along NMDS2. For the four inner harbour sites, positive NMDS2 scores were associated with dominance of *C. rhizophorae* in death assemblages while negative NMDS2 scores were associated with dominance of *I. alatus*. Except for GSP, compositional differences between living and dead were significant for all sites (PERMANOVA; $p < 0.05$).

Multivariate dispersion varied across Kingston Harbour with the lowest apparent values at GSP and highest at BS (Figure 4.7). Visually, death assemblages were consistently less dispersed at each site compared to the living community. However, live-dead differences were only significant at BS, GC, and RC (Randomization test; $p < 0.05$). The low number of samples at GSP and OCW likely contributed to the insignificant test results therein despite visual separation.

4.4.4 Vertical positioning of dominant taxa

While not formally quantified, we noted during field sampling that subtidal sections of mangrove roots and wharf surfaces were consistently and predominantly occupied by dense clusters of *I. alatus* while a more even mix of all species was found to occupy the intertidal zone. Compared with images during the peak of the *P. viridis* invasion by Buddo (2008) and our finding that *C. rhizophorae* dominates death assemblages, we reconstruct a hypothesized history of Kingston Harbour's epibiont bivalve community as having experienced three phases (Figure 4.8): (1) a pre-invasion assemblage dominated by *C. rhizophorae*, (2) an invasion assemblage dominated by *P. viridis*, resulting in the new species also appearing in the death assemblage, and (3) a post-invasion assemblage dominated by *I. alatus*, but also featuring the more recent invader *M. strigata*.

4.5 Discussion

4.5.1 Green mussels: death assemblages remember the intensity of past invasions

We found that the densities of *P. viridis* within death assemblages across Kingston Harbour were, remarkably, perfectly correlated to those reported living during the early peak of the invasion (Figure 4.3). The ability of death assemblages to capture this spatial variation is indicative that they may serve as an effective proxy with which to retroactively reconstruct the early invasion history at a broader scale across the harbour and elsewhere in the west Atlantic. Although the same correlation holds for the current living population 20 years following introduction, the extreme rarity of *P. viridis* today (only 15 individuals among our 35 samples) makes this correlation less convincing and, possibly, entirely a product of chance. A far more rigorous survey of the standing *P. viridis* population in Kingston Harbour is required to detect patterns of spatial variation today.

Buddo et al. (2003) reported mean *P. viridis* densities as high as 1,539.20 individuals/m^2, and for a single monthly sample, >8,000 individuals/m^2. Despite how densely populated *P. viridis* was at the time, it represented the least abundant species in both the current living community and the death assemblage, suggesting that (1) a single large pulse of shell production generated most dead shells, (2) taphonomic bias may be acting on these shells, and/or (3) human efforts to remove living *P. viridis* prevented shells from incorporation in death assemblages. A single production pulse is supported by monthly densities reported by Buddo et al. (2003) as most monitored sites experienced a short-term peak in relative densities. The complete lack of dead shells at GSP, which had the lowest densities during the peak of the invasion (Figure 4.3), further supports that the majority of dead *P. viridis* shells derived from a single production pulse with minimal continued inputs, resulting in the species being relatively rare in the death assemblage after the initially rapid destruction of

most shells (Tomašových et al., 2016). On the other hand, the majority of *P. viridis* shells that we recovered were heavily fragmented, exhibited signatures of predation (boreholes, etc.) and had undergone chemical alteration (loss of colouration and periostracum, chalky interiors, etc.). *Perna viridis* was harvested by artisanal fisheries for human consumption and fishing bait early into the invasion history, but the detection of numerous harmful biological, chemical, and metallic contaminants in the tissues of wild populations dissuaded widespread harvest (Buddo et al., 2012). Human intervention is unlikely to have sufficiently contributed to population decline due to a lack of dedicated culling efforts. Despite these limits on the quantity of shell material – all of which are worth further investigation – densities that have survived appear to strongly predict the relative size of known historic populations.

What this study does not reveal are insights into the potential causes of the harbour-wide boom-bust dynamics of *P. viridis*, which remain speculative. Strayer et al. (2017) outlined several plausible mechanisms through which boom-bust invasions may occur, including (1) accumulation of enemies (new predators, parasites, etc.), (2) density-dependent time-lags, (3) delayed genetic effects, (4) interactions with subsequent invaders, and (5) human control. The first three mechanisms are all plausible and may not be mutually exclusive. Regarding enemies (mechanism 1), durophagous predators are abundant in Kingston Harbour, including decapods and drilling gastropods like *Murex sp.* (Alleng, 1997). In a series of controlled predation experiments, Mitchem et al. (2007) found *P. viridis* to be preyed upon by blue crabs (*Callinectes sapidus*) at twice the rate of oysters (*Crassostrea virginica*) and quahogs (*Mercenaria mercenaria*) but were not consumed by Caribbean spiny lobsters (*Panulirus argus*). As stated above, the prevalence of fractured and drilled dead shells indicates that local predators preyed upon *P. viridis* following their rise to dominance. Density-dependent time-lags (mechanism 2) likely contributes due to the fast rates at which *P. viridis* both grows and colonizes new surfaces (Buddo, 2008; Elliott et al., 2012). Once the carrying capacity is met or surpassed, failure for future generations to obtain optimal space for fouling and growth

may cause the population to subsequently crash. Delayed genetic effects (mechanism 3) is particularly supported by the low genetic variation among *P. viridis* populations introduced throughout the west Atlantic (Gobin et al., 2013; Gilg et al., 2014). Inbreeding depression may have debilitated the population's ability to resist diseases or parasites, tolerate Kingston Harbour's stressful environmental conditions (Wade et al., 1972; Rose and Webber, 2019), or maintain high reproductive output. Interactions with subsequent invaders (mechanism 4) is unlikely as no subsequent biofouling invaders have been documented until the present study (but see below). For similar reasons highlighted above regarding human removal (mechanism 5), deliberate intervention efforts were likely insufficient.

4.5.2 Native bivalves: the unpredictability of post-invasion ecological recovery

The decline of *P. viridis* from the subtidal sections of mangrove roots and wharf surfaces in Kingston Harbour allowed the native *I. alatus* to reclaim and dominate these surfaces. However, death assemblage compositions suggested that the community was historically dominated by *C. rhizophorae* (Figures 4.4-4.5). Live-dead discordance may have numerous, non-mutually exclusive explanations that should be evaluated in order of parsimony (e.g., the five hypotheses from Bizjack et al., 2017). Here, shortcomings in sample collection or processing (hypothesis 1) and contamination due to the dumping of dredge spoils (hypothesis 2) are rejected as we believe our sampling procedures were consistent, we processed every shell in each sample, and the shallow mangrove areas of Kingston Harbour are far from dredge dump sites.

The effects of time averaging (hypothesis 3) are several-fold. First, it is possible that the window of time-averaging in these death assemblages is long enough to capture, but obscure, numerous community turnover events wherein I. alatus or other taxa were dominant at points in the past, reinforcing a false sense of stability (Tomašových and Kidwell, 2010). Post-mortem shell dates of both *C. rhizophorae* and *I. alatus*, obtained via radiocar-

bon or amino acid racemization techniques, would aid in determining whether differential production among taxa has occurred over the window of time-averaging. Second, spatial heterogeneity across Kingston Harbour in the past may have been smeared as community structure changed over time (e.g., Kidwell and Tomašových, 2013; Tyler and Kowalewski, 2017; but see Tomašových et al., 2020). In Kingston Harbour, death assemblages exhibited reduced variation compared to living communities, but this was only detected statistically at three of our five sites (Figures 4.6-4.7).

Preservational bias (hypothesis 4) is potentially an important factor as the two dominant species – *C. rhizophorae* and *I. alatus* – notably differ in characteristics thought to affect preservation potential (Table 4.2). The thin, aragonitic, and organic-rich shell of *I. alatus* suggests that this taxon is far more susceptible to post-mortem destruction, and many dead individuals were indeed represented by isolated, but distinctive, fragments possessing hinges (although, so too were those of *C. rhizophorae*). However, skeletal elements of dissolution-prone taxa within Bivalvia, and even across other phyla, have shown promise for surviving on similar timescales within sediment records (Nawrot et al., 2022). In contrast, the cementing behaviour of *C. rhizophorae* can conceivably result in underrepresentation in the death assemblage as the cemented valve remains attached to the substrate while the other valve falls into the sediment. Our *C. rhizophorae* shells did not appear to be biased towards one valve or the other and very few (or no) dead shells remained attached to mangroves at the stations we sampled. Even if we were to divide the number of *C. rhizophorae* shells in half to account for this effect, they would remain the dominant taxon observed in the death assemblage.

Genuine ecological change (hypothesis 5) is also supported. Siung (1980) previously noted that over-exploitation of *C. rhizophorae* in Jamaica was detectable and proposed *I. alatus* as a potential alternative for commercial harvest. Nonetheless, based on settling rates on artificial substrates, Elliott et al. (2012) found *C. rhizophorae* to still be an abundant

and effective biofouler decades later. Following a juvenile phase of byssal-attachment, *C. rhizophorae* cements itself to its substrate and requires a large attachment surface, while the life-long byssal-attaching *I. alatus* can efficiently occupy small surfaces without hindering development and growth (Bromley and Heinberg, 2006). Thus, the subtidal takeover of *I. alatus* we observed in 2019 perhaps occurred opportunistically and concurrently to the decline of *P. viridis*. Other confounding factors, such as differential tolerances to the conditions of Kingston Harbour today, may have contributed to the disproportionally large recovery of *I. alatus* compared to *C. rhizophorae*, but these factors cannot be elucidated here. It is noteworthy, however, that the proportional abundance of *B. exustus* remained constant between the living community and death assemblages, suggesting that recovery was taxon-specific.

As many high-profile invasion events continue to harm biological communities, management decisions may be understandably focused on removal efforts without equal emphasis on subsequent ecological recovery. Live-dead discordance in Kingston Harbour, driven primarily by the dramatic reversal in dominance of two native taxa, suggests that the community has fallen into a novel community configuration. Brief (and largely unsuccessful) invasion events like that of *P. viridis* examined here are powerful tests by which the nature of ecological recovery can be explored. In a meta-analysis of 151 such case studies, Prior et al. (2018) found that 31% of cases resulted in negative recovery outcomes following the removal of an invasive species and a mix of positive and neutral outcomes for the remainder. The outcome in Kingston Harbour is thus not unexpected but will require additional conservation measures if the recovery of the economically-important *C. rhizophorae* to historic population densities is desired.

4.5.3 Charu mussels: death assemblages lag behind recent invasion events

To our knowledge, this study represents the first reporting of *M. strigata* in Jamaica. Native to the Pacific and Atlantic coasts of South America as well as the eastern Pacific as far north as Mexico, *M. strigata* was first introduced to Florida as early as 1986, although considered extirpated after only a few years until reappearing in 2004 (Boudreaux and Walters, 2006). More recently, *M. strigata* has been introduced to Southeast Asia with reported sightings from India (Jayachandran et al., 2019), Singapore (Lim et al., 2018), the Philippines (Rice et al., 2016), and Taiwan (Huang et al., 2021). It has been suggested that these trans-hemisphere introductions may considerably harm fisheries of native *P. viridis* in Southeast Asia (Sanpanich and Wells, 2019). The rapid introduction of *M. strigata* to Southeast Asia is important to note as it may be related to our discovery in Jamaica; the taxon is currently spreading elsewhere in the Caribbean, as recently reported for previously uninhabited areas of Venezuela (Lodeiros et al., 2021). *Mytella strigata* may thus be a particularly effective invader that only recently has been given the opportunity of transplanting throughout the tropics.

The exact timing of the introduction of *M. strigata* to Kingston Harbour is unclear, but the absence of historic and modern records suggest it is recent (<10 years) and under-representation of *M. strigata* in our death assemblages supports this timing. Taphonomic processes are unlikely to destroy *M. strigata* shells at a particularly high rate compared to other taxa as they should have comparable durability as *I. alatus* and potentially greater durability than *B. exustus* (Table 4.2), both of which are more common in the death assemblage (Figure 4.4). The introduction is also likely to have taken place several years ago as there was must have been sufficient time for adult populations to establish in the living community and shells from at least the first generation to accumulate in the underlying death assemblage. Unlike *P. viridis*, *M. strigata* appears to have not undergone a notable early population boom in Kingston Harbour, but rather, has gradually incorporated itself into

the community. *Mytella strigata* may still be in an initial lag phase prior to an impending boom (Strayer et al., 2017) or has failed to boom, possibly due to the nature of ecological recovery following the decline of *P. viridis* (e.g., competition with *I. alatus*). Gradual population growth would also explain how *M. strigata* has a larger share of death assemblage occupation than *P. viridis*: constant shell production would continually supply the death assemblage even under conditions favouring high disintegration rates versus a single, large pulse of input (Tomašových et al., 2016). As the window of time-averaging for *M. strigata* in the death assemblage is constrained by the time of introduction, the rapid accumulation of shells thus far is a testament to how fast the signal of a new taxon may be detected in fossil assemblages, even if the remains of established taxa are coarsened to orders of magnitude greater intervals.

4.6 Conclusions

Using living communities and death assemblages of epibiont bivalves from mangrove habitats in Kingston Harbour, Jamaica, we found that (1) spatial variation of the boom-bust invasion of Asian green mussels (*Perna viridis*) at the turn of the century was preserved in death assemblages, making the dead an effective proxy of invasion history, (2) taxonomic structure of the bivalve community was reconfigured following the peak of the invasion such that the dominant and economically-valuable mangrove cup oyster (*Crassostrea rhizophorae*) was largely replaced by the flat tree oyster (*Isognomon alatus*), likely due to opportunistic colonization of subtidal surfaces previously occupied by *P. viridis*, and (3) the overrepresentation of the previously unreported charru mussel (*Mytella strigata*) in the living community, signalling that this taxon was introduced more recently to Kingston Harbour. Our findings exemplify the under-exploited value of geohistorical data to reveal insights on biological invasions and their lasting impacts on ecological recovery. Such insights are particularly valuable for areas where ecological evaluation is urgently required, yet resources for long-term monitoring

efforts are scarce.

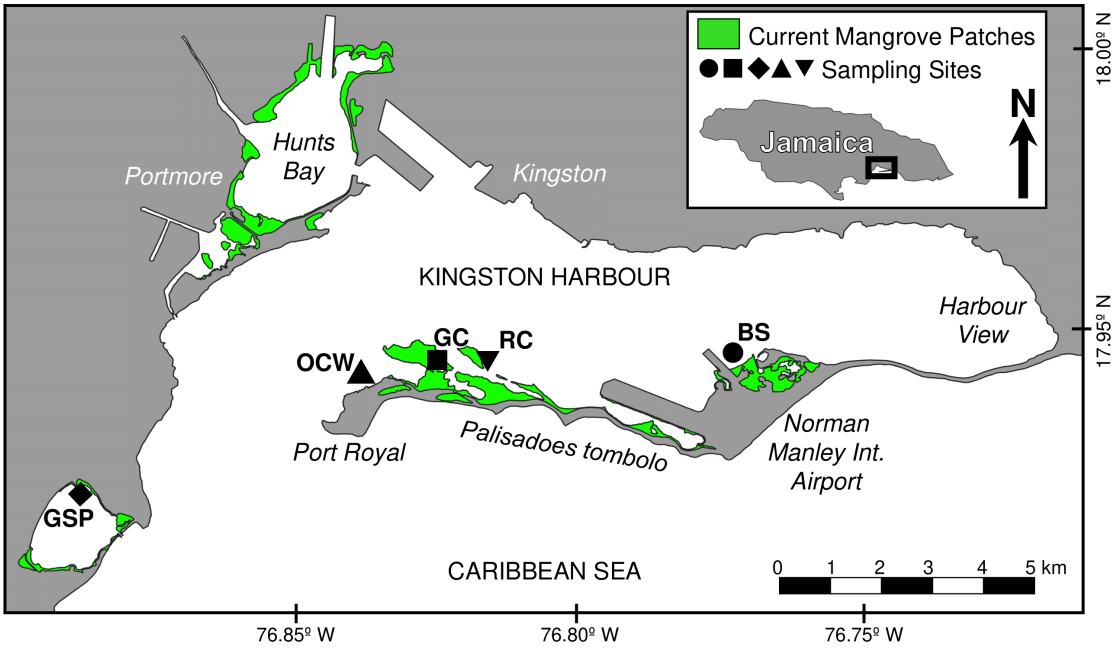

Figure 4.1: Map of the Kingston Harbour on the south coast of Jamaica, showing the locations of sites from which living and dead bivalves were sampled and the current extents of large mangrove patches. The five sampling sites coincide with areas from which *Perna viridis* population densities were previously monitored at the turn of the century (Buddo et al., 2003). Site abbreviations: BS, Buccaneer Swamp; GC, Goodbody Channel; GSP, Great Salt Pond; OCW, Old Coal Wharf; RC, Refuge Cay.

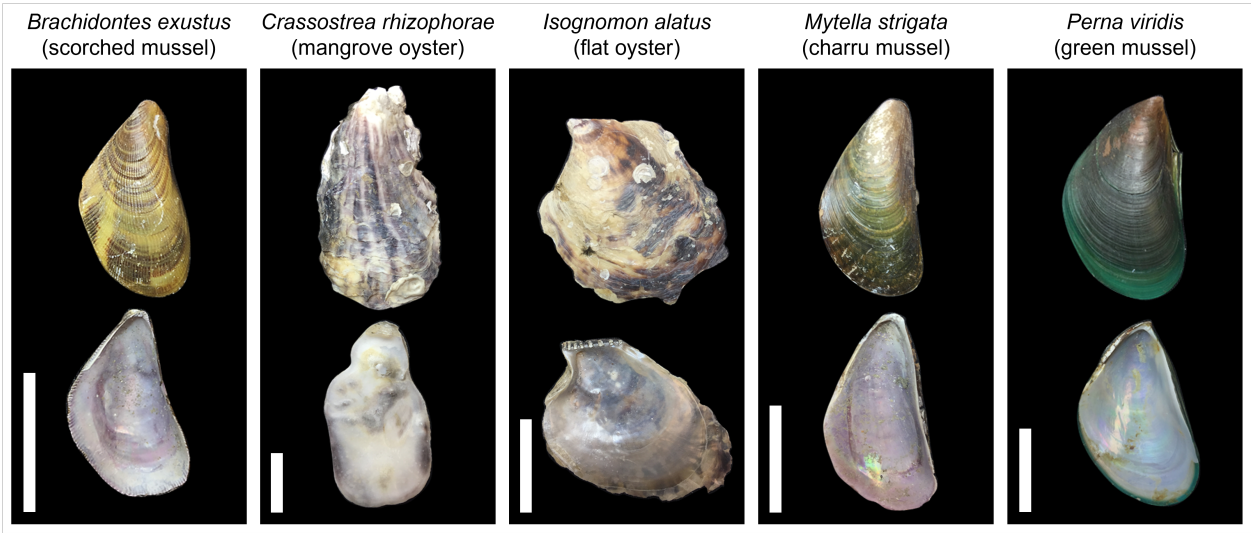

Figure 4.2: Representative dead-collected shells of the five mangrove-dwelling bivalve species assessed in this study, showing both valves from originally-articulated specimens. From left to right: *Brachidontes exustus* (scorched mussel), *Crassostrea rhizophorae* (mangrove oyster), *Isognomon alatus* (flat tree oyster), *Mytella strigata* (charru mussel), *Perna viridis* (Asian green mussel). Scale bars = 10 mm.

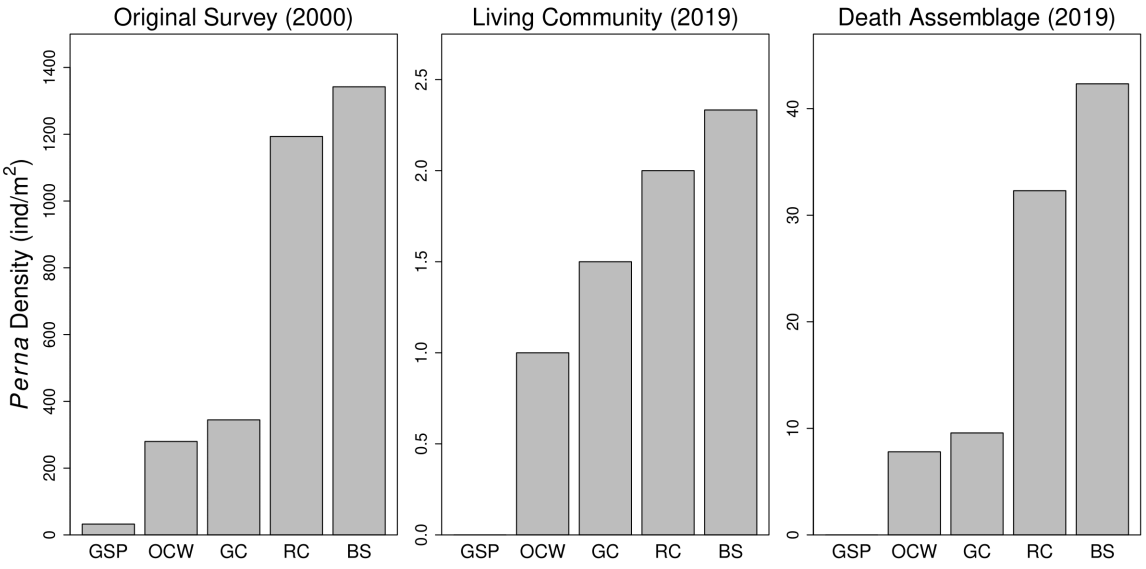

Figure 4.3: Mean densities of Perna viridis from three datasets. Left: living *P. viridis* reported by Buddo et al. (2003) and Buddo (2008). Middle: current living individuals surveyed in 2019. Right: dead shells collected in 2019. No *P. viridis* specimens were observed currently living or dead at GSP in 2019. Site abbreviations: BS, Buccaneer Swamp; GC, Goodbody Channel; GSP, Great Salt Pond; OCW, Old Coal Wharf; RC, Refuge Cay.

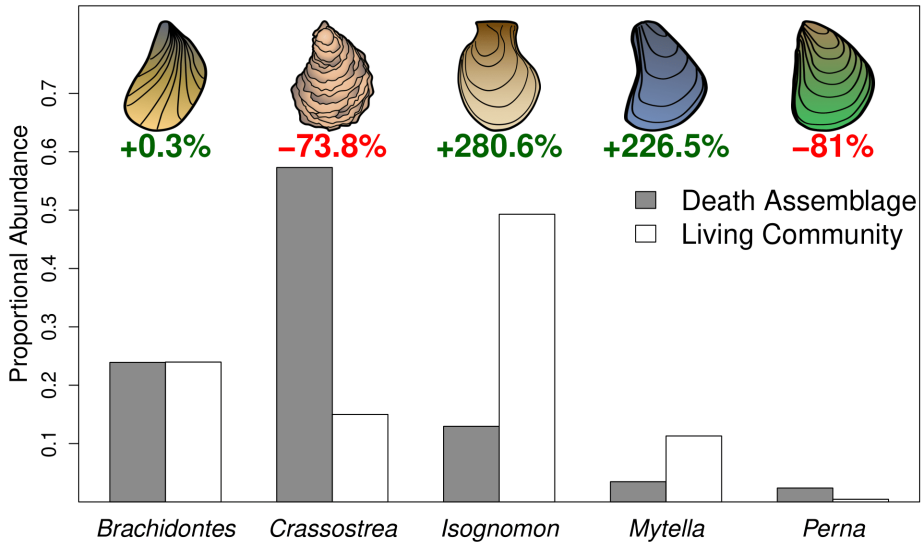

Figure 4.4: Proportional abundances of mangrove-dwelling bivalves in the living community and death assemblage summed across all sites. Percent increases or decreases for each species are printed above paired bars. *Brachidontes exustus* abundances are nearly identical between the death assemblage and living community. *Crassostrea rhizophorae* is the dominant species found in the death assemblage, but *Isognomon alatus* is dominant in the living community. *Mytella strigata*, the more recent of the two invaders, is disproportionately rare in the death assemblage. *Perna viridis*, the historic invader, is conversely overrepresented in the death assemblage, reflecting the short-lived nature of their dominance at the turn of the century. The overall live-dead discordance remains moderately high as only *C. rhizophorae* and *I. alatus* switch places in ranked order.

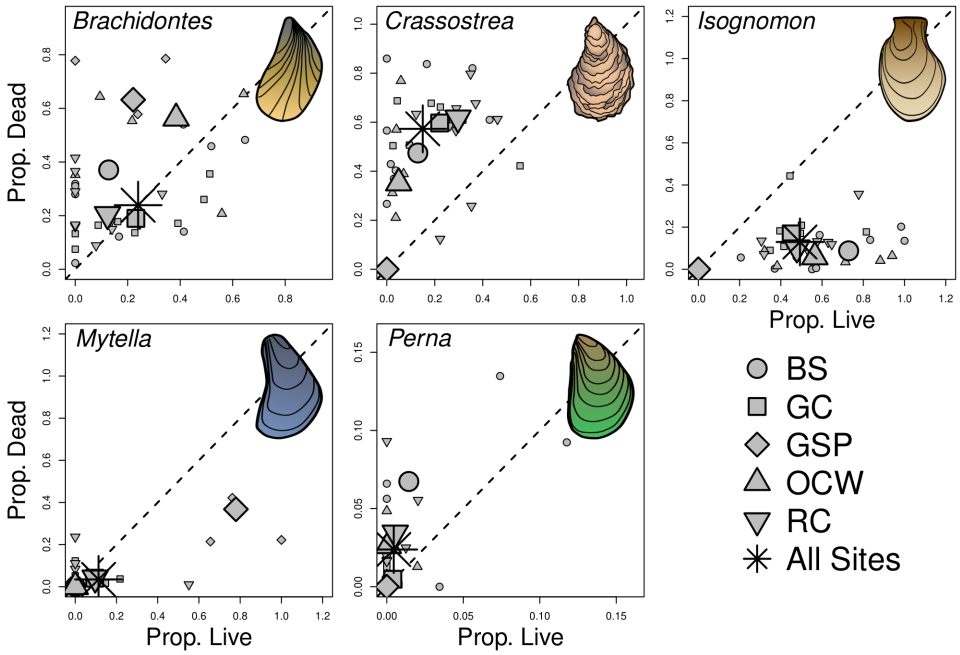

Figure 4.5: Live versus dead proportional abundances at per-sample (35; small icons), per-site (5; large icons) and summed total (star icons) for each species. Samples falling on either side of the dashed line indicates overrepresentation in the death assemblage (above the dashed line) or the living community (below the dashed line). *Brachidontes exustus* proportional abundances are generally higher in the dead using sample- and site-level data despite being equal for the summed data. Summed proportional abundance trends for the remaining four taxa are similarly reflected by sample- and site-level data such that *Crassostrea rhizophorae* and *Perna viridis* are overrepresented in the death assemblage, while *Isognomon alatus* and *Mytella strigata* are overrepresented in the living community. Site abbreviations: BS, Buccaneer Swamp; GC, Goodbody Channel; GSP, Great Salt Pond; OCW, Old Coal Wharf; RC, Refuge Cay.

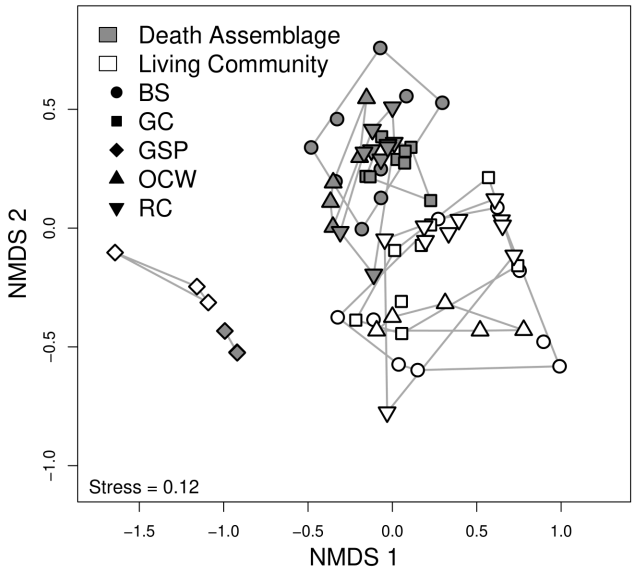

Figure 4.6: Non-metric multidimensional scaling (NMDS) ordination of Bray-Curtis dissimilarities based on Hellinger-transformed abundance data from living and death assemblage samples. Convex hulls indicate score ranges for assemblages from each of the four sites. Living and death assemblages are compositionally distinct in ordination space, but the four sites within the inner margin of Kingston Harbour (BS, GC, OCW, and RC) are not strongly separated for either assemblage type. Site abbreviations: BS, Buccaneer Swamp; GC, Goodbody Channel; GSP, Great Salt Pond; OCW, Old Coal Wharf; RC, Refuge Cay.

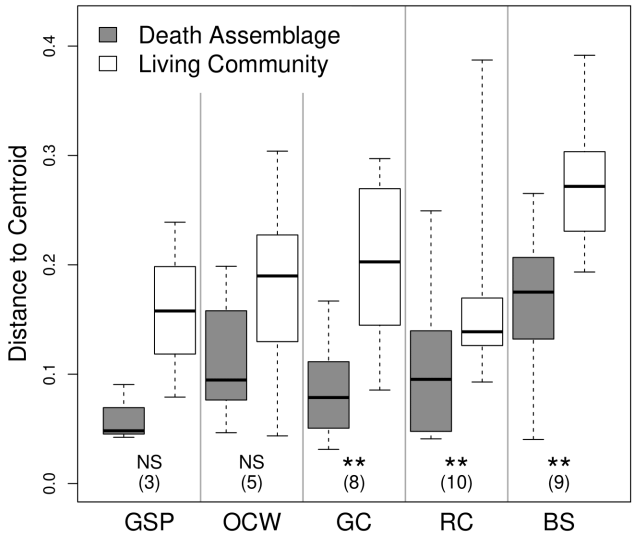

Figure 4.7: Boxplots for the dispersion of samples around multivariate group centroids (i.e., within-assemblage compositional variation). The living community exhibited significantly higher group dispersion than the corresponding death assemblage at GC, RC, and BS (Randomization test, $p < 0.05$), which may be a product of spatial homogenization due to time-averaging in the death assemblage or a signal of increased compositional heterogeneity in the living community. Parenthetical values indicate the number of live-dead sample pairs per group. Note that despite visual separation of boxes at GSP and OCW, low sample sizes did not provide adequate power to test for significant differences at these sites. Site abbreviations: BS, Buccaneer Swamp; GC, Goodbody Channel; GSP, Great Salt Pond; OCW, Old Coal Wharf; RC, Refuge Cay.

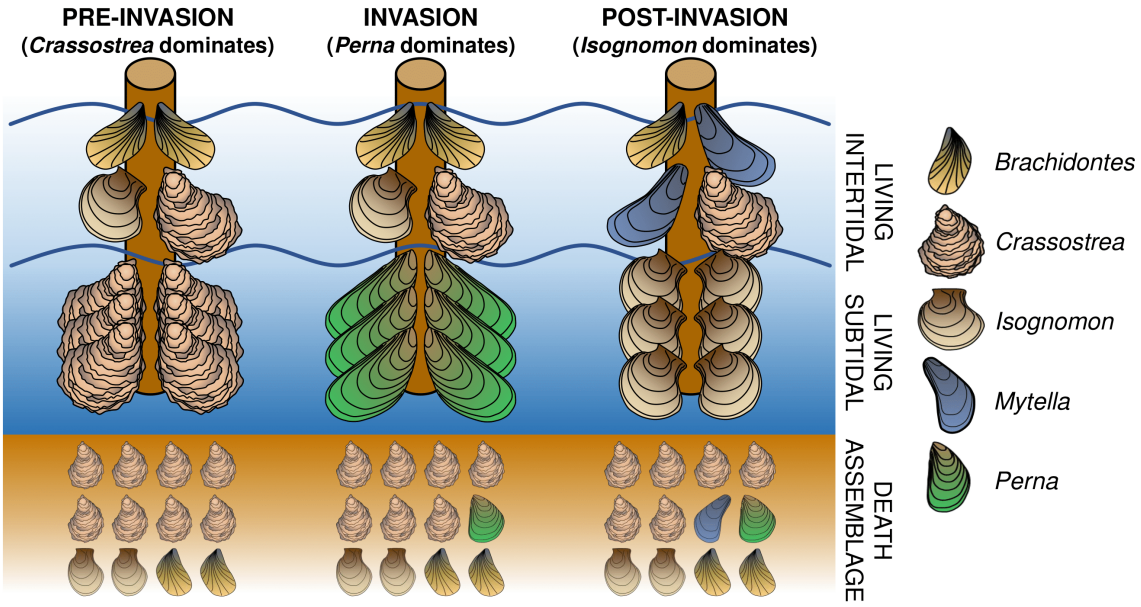

Figure 4.8: Illustrative summary of changing mangrove communities in Kingston Harbour based on field observations in 2019. Left: death assemblages indicated that prior to the arrival of *Perna viridis* in 1998, *Crassostrea rhizophorae* was the dominant epibiont bivalve living on mangrove roots and wharf pilings. Middle: shortly after initial detection in 1998, *P. viridis* rapidly overtook subtidal surfaces and dominated for several years. *Perna viridis* shells were incorporated into the death assemblage, leaving a record from which past population densities may be estimated. Right: living communities from 2019 indicated that the native fauna, occupying the intertidal zone, recolonized subtidal surfaces as *P. viridis* populations declined by the 2010s. However, poor live-dead fidelity suggested that the succeeding community fell into a novel configuration where *Isognomon alatus* is now the dominant taxon, particularly in the subtidal zone. The previously undocumented presence of *Mytella strigata*, coupled with its disproportionately low representation in death assemblages, suggested that this species is a much more recent invader.

Table 4.1: Summary of field sites from which samples were collected.

Site Name	Abbrev.	Lat.	Long.	Fouling Surface	N Samples
Buccaneer Swamp	BS	17.94381	-76.77752	Mangroves/Warfs	9
Goodbody Channel	GC	17.94338	-76.82746	Mangroves	8
Great Salt Pond	GSP	17.92153	-76.89278	Mangroves	3
Old Coal Wharf	OCW	17.94140	-76.837645	Warfs	5
Refuge Cay	RC	17.94468	-76.82156	Mangroves	10

Table 4.2: Counted abundances and shell characteristics of the five mangrove-dwelling bivalves.

	Brachidontes exustus	*Crassostrea rhizophorae*	*Isognomon alatus*	*Mytella strigata*	*Perna viridis*
Dead	6,857	16,437	3,715	993	683
Live	793	496	1,631	374	15
Dead + Live	7,650	16,933	5,346	1,367	698
Status	Native	Native	Native	Introduced	Introduced
Attachment	Byssus	Cementation	Byssus	Byssus	Byssus
Mineralogy	Aragonite	Calcite	Bimineralic	Aragonite	Aragonite
Max Size*	Small	Large	Medium	Medium	Large
Thickness**	Thin	Thick	Thin	Thin	Thick
Organics***	High	Low	High	High	High

*Typical reported max shell length: Small (<5 cm), Medium (5-10 cm), Large (>10 cm)
**Shell thickness: Thin (<0.5 mm), Thick (>0.5 mm)
***Density of organic carbon within shell matrix: Low (<0.5 wt%), High (>0.5 wt%)

APPENDIX A

ABBREVIATIONS

Table A.1: Abbreviations used throughout the text.

Abbreviation	Meaning
ABC	Abundance/Biomass Comparison
AMBI	ATZI's Marine Benthic Index
ANOSIM	Analysis of Similarity
BACI	Before-after-control-impact
BENTIX	Benthic Index
BRI	Benthic Response Index
DDT	Dichlorodiphenyltrichloroethane
ECY	Washington State Department of Ecology
EG	Ecological Group
EQS	Ecological Quality Status
IQR	Interquartile Range
JWPCP	Joint Water Pollution Control Plant
LACSD	Los Angeles County Sanitation Districts
M-AMBI	Multivariate ATZI's Marine Benthic Index
NMDS	Non-metric Multidimensional Scaling
NOAA	National Oceanic and Atmospheric Administration
NSF	National Science Foundation
OCSD	Orange County Sanitation Districts
OTU	Operational Taxonomic Unit
P/A	Presence/Absence
PCB	polychlorinated biphenyl
PERMANOVA	Permutational Multivariate Analysis of Variance
PIE	Probability of Interspecific Encounter
PRF	Petroleum Research Fund
PSEMP	Puget Sound Ecosystem Monitoring Program
PSEP	Puget Sound Estuary Program
SAB	Species-abundance-biomass
SCAMIT	Southern California Association of Marine Invertebrate Taxonomists
SIMPER	Similarity Percentage Procedure

APPENDIX B

CHAPTER 1 DATA AND CODE

The full dataset and R code for Chapter 1 is available in the online supplementary materials under the following folder:

Kokesh2024_ Appendix_ B_ Chapter_ 1.zip.

APPENDIX C

CHAPTER 2 DATA AND CODE

The full dataset and R code for Chapter 2 is available in the online supplementary materials under the following folder:

Kokesh2024_ Appendix_ C_ Chapter_ 2.zip.

APPENDIX D

CHAPTER 3 DATA AND CODE

The full dataset and R code for Chapter 3 is available in the online supplementary materials under the following folder:

Kokesh2024_ Appendix_ D_ Chapter_ 3.zip.

APPENDIX E

CHAPTER 4 DATA AND CODE

The full dataset and R code for Chapter 4 is available in the online supplementary materials under the following folder:

Kokesh2024_ Appendix_ E_ Chapter_ 4.zip.

APPENDIX F

ADDITIONAL ANALYSES OF BIOTIC INDICES

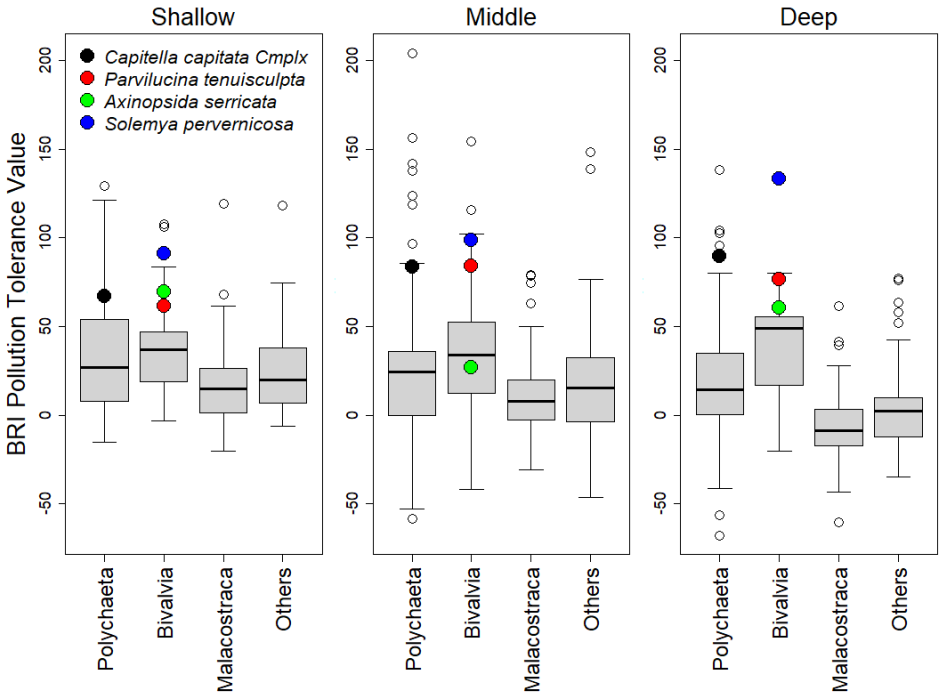

Figure F.1: Boxplots of pollution-tolerance values assigned to species encountered on the Palos Verdes shelf, with the three most abundant classes displayed separately. Tolerance values were assigned by Smith et al. (2001) based upon the position of a taxon's highest abundance along a known pollution gradient, which had been quantified by ordination of sediment chemistry (see methods in Chapter 3). These tolerance scores are weighted by taxonomic abundances to calculate the BRI of a sample. The mean tolerance values for bivalves were found to be higher than for any other group, and thus bivalves can be expected to yield generally higher (i.e., overestimated) BRI scores than would be calculated using the whole fauna. The tolerance values of four species are highlighted by colored points: the famously pollution-tolerant polychaete *Capitella capitata Cmplx* (black), the facultative chemosymbiotic bivalves *Parvilucina tenuisculpta* (red) and *Axinopsida serricata* (green), and the obligate chemosymbiontic bivalve *Solemya pervernicosa* (blue). Pollution values assigned to these taxa mostly fall above the IQR of their clade's distribution, i.e. toward the pollution-tolerant end of the spectrum. In the shallow- and middle-depth species lists, the two bivalve taxa having the highest pollution tolerance values (outlier points) were the mixed suspension-deposit-feeding *Macoma nasuta* and *M. carlottensis*.

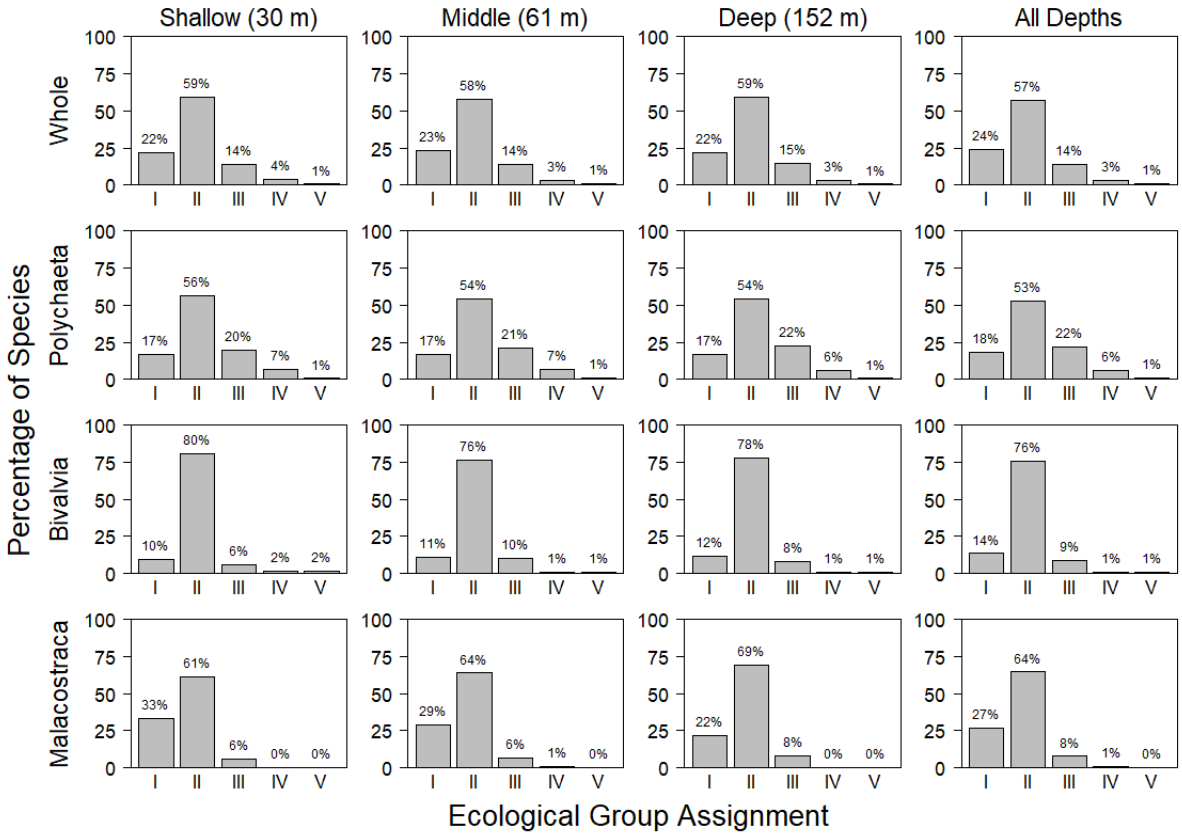

Figure F.2: Bar plots of the percentage of species within taxonomic subsets assigned to each ecological group (EG) for the calculation of AMBI scores. In all cases, the majority of taxa on the US-AMBI species list were assigned to EGII (top row). Polychaetes appeared to be the best match of the distributions seen in the whole fauna, and thus might be expected to serve as the strongest surrogate for calculating AMBI. In contrast, the distributions for bivalves and malacostracans skewed more heavily towards EGI and EGII. They are thus expected to underestimate AMBI scores calculated from the whole fauna. For bivalves, the only species that occupy EGIV and EGV (characterized as second- and first-order opportunists) are *Parvilucina tenuisculpta* and *Solemya pervernicosa*, respectively.

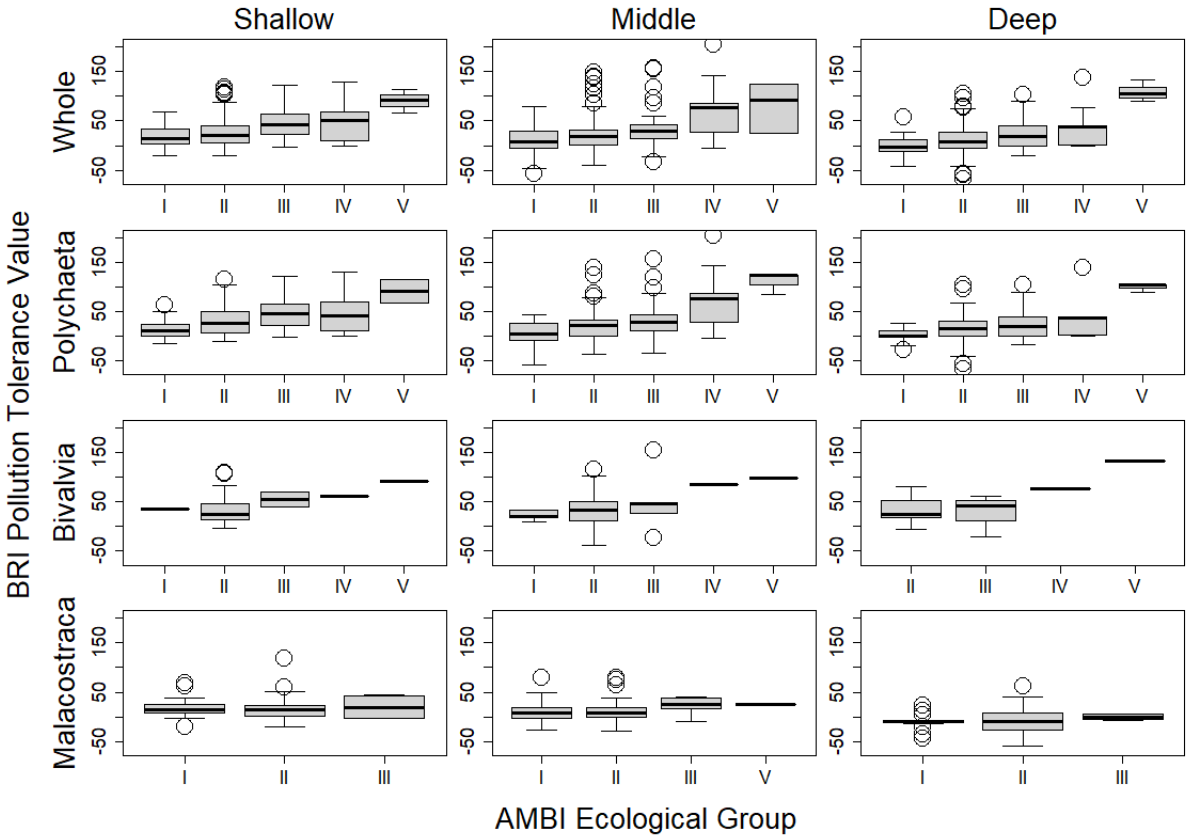

Figure F.3: The relationship between BRI pollution tolerance values and AMBI ecological groups (EGs) assigned to each species on the Palos Verdes shelf. In general, taxa with high pollution tolerance values tend to also be assigned to more tolerant EGs. However, the variation of pollution tolerance values within each EG is greater than most differences in mean among EGs. Among bivalves (third row), only two or at most three EG categories include sufficient taxa to support a "box", greatly reducing the potential power of this class.

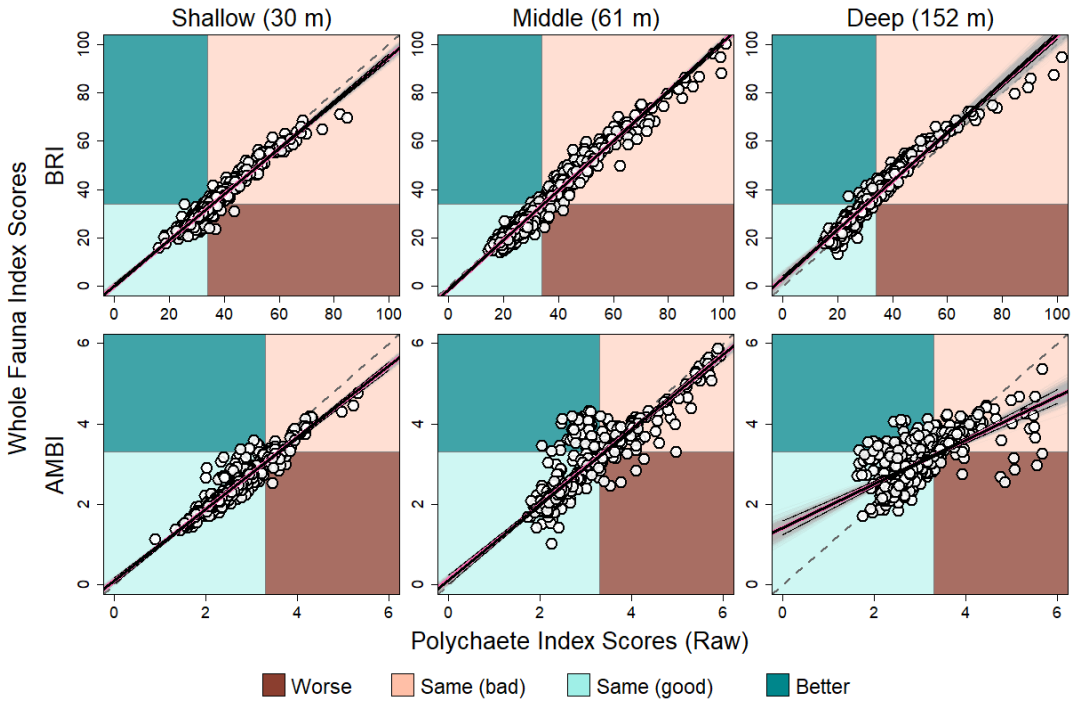

Figure F.4: BRI (top row) and AMBI (bottom row) scores generated using the whole fauna versus scores generated using only polychaetes. Analyses are a replication of those from Chapter 3 with polychaetes substituted for bivalves (see Figure 3.3 for details). For both BRI and AMBI, polychaetes exhibited particularly strong correlation with scores generated using the whole fauna, and most points fall into the two 'Same' status quadrants, indicating high agreement in ecological quality status (EQS). This high correlation – far stronger than seen with bivalves – is likely due to polychaetes representing (1) the majority of all individuals in the whole fauna and (2) having a fairly strong representation of all tolerance categories, very similar to that of the whole fauna (Fig. F2), making this class a particularly strong surrogate for estimating the EQS of the whole fauna.

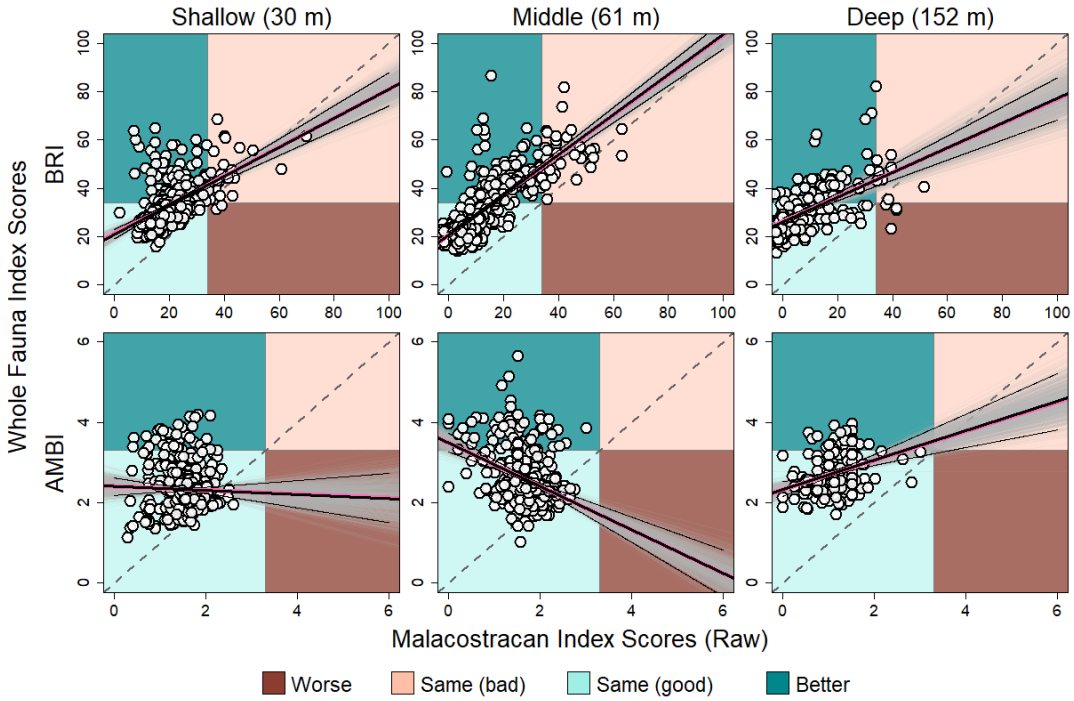

Figure F.5: BRI (top row) and AMBI (bottom row) scores generated using the whole fauna versus scores generated using only malacostracans. Analyses are a replication of those from Chapter 3 with malacostracans substituted for bivalves (see Figure 3.3 for details). BRI scores exhibited positive correlations for all isobaths, but malacostracan-BRI generally underestimated whole-BRI with many points occupying the top-left 'Better' status quadrant (e.g., malacostracans incorrectly indicated 'Good' status when whole-BRI would indicate 'Poor' status). Malacostracan-AMBI similarly underestimated whole-AMBI. However, given the variable slopes and wide bootstrapped confidence intervals, these samples demonstrated only a weak correlation with whole-AMBI values.

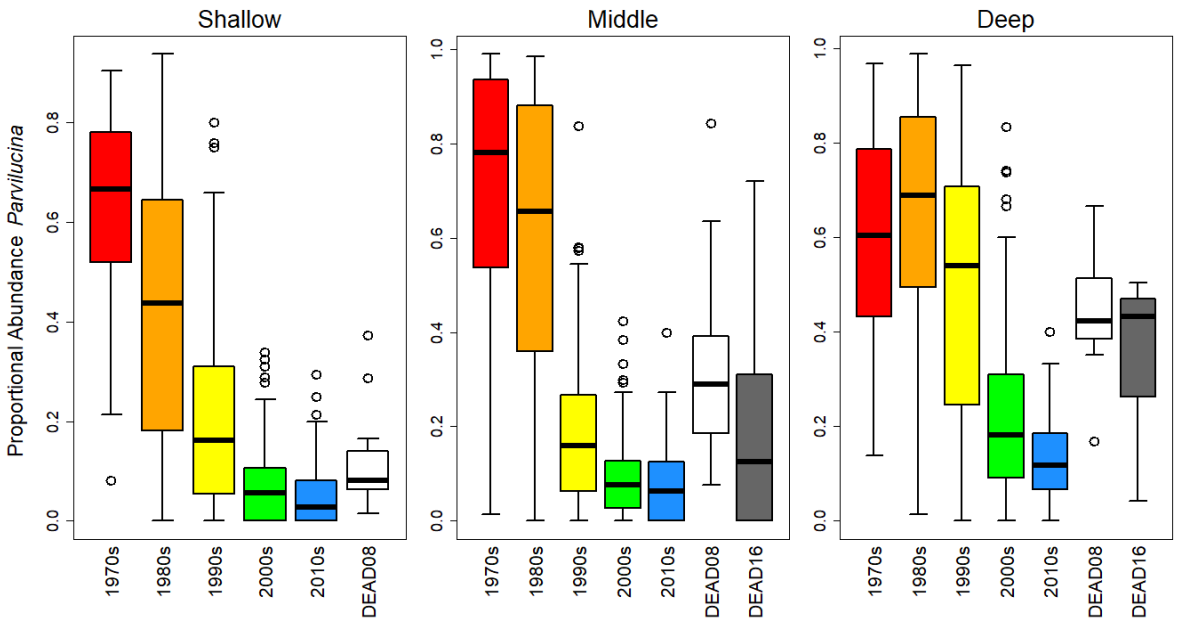

Figure F.6: Boxplot of the proportional abundances of *Parvilucina tenuisculpta* of bivalve living and death assemblages. *Parvilucina tenuisculpta* was the dominant bivalve species at most sites during the 1970s-1980s before its sharp decline during the 1990s; it persisted in low abundances up to the 2010s. However, death assemblages collected in 2008 and 2016 both exhibited mean proportional abundances of *P. tenuisculpta* that were significantly higher than their corresponding living assemblages in those decades (2000s and 2010s) based on the IQR of each decade's annually-collected samples. The proportional abundance of this taxon in death assemblages generally increased with depth in each decade, whereas its abundance in the living assemblages remained relatively consistent across depth (except for the 1990s). *Parvilucina tenuisculpta* is thus detectably overrepresented in death assemblages collected in 2008 and 2016, likely because they retain shells from the high input of large-bodied individuals that occurred during the population boom in preceding decades.

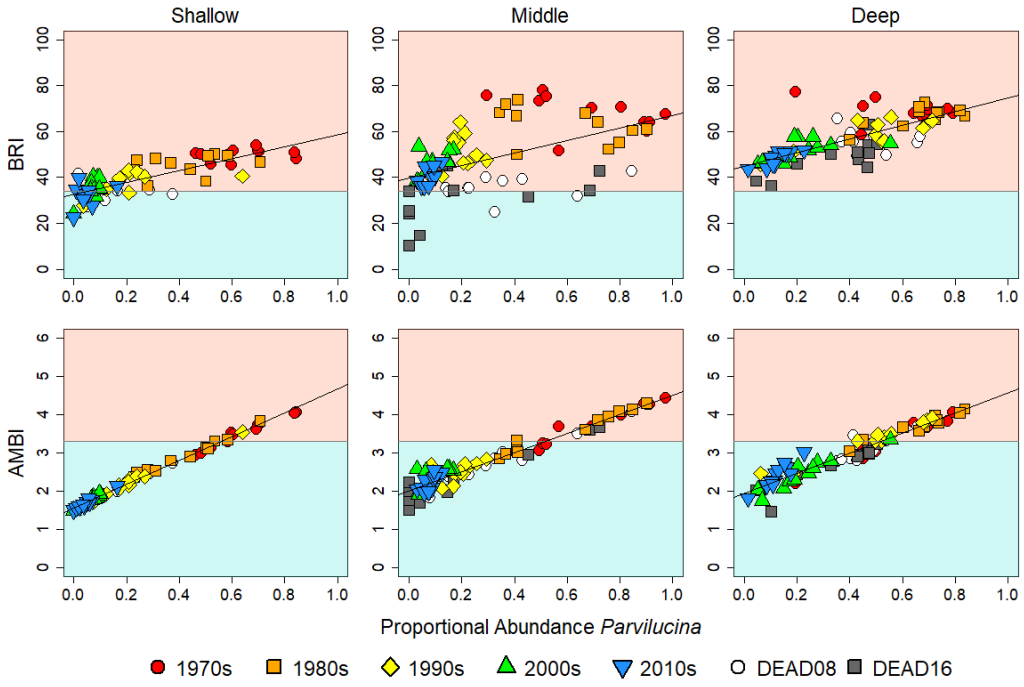

Figure F.7: The relationship between mean bivalve-BRI (top row) and bivalve-AMBI (bottom row) versus the mean proportional abundance of *Parvilucina tenuisculpta* at each station and during each decade of live-collected data, as well as the death assemblages collected in 2008 and 2016. Proportional abundances of *P. tenuisculpta* were positively correlated to both BRI and (especially) AMBI.

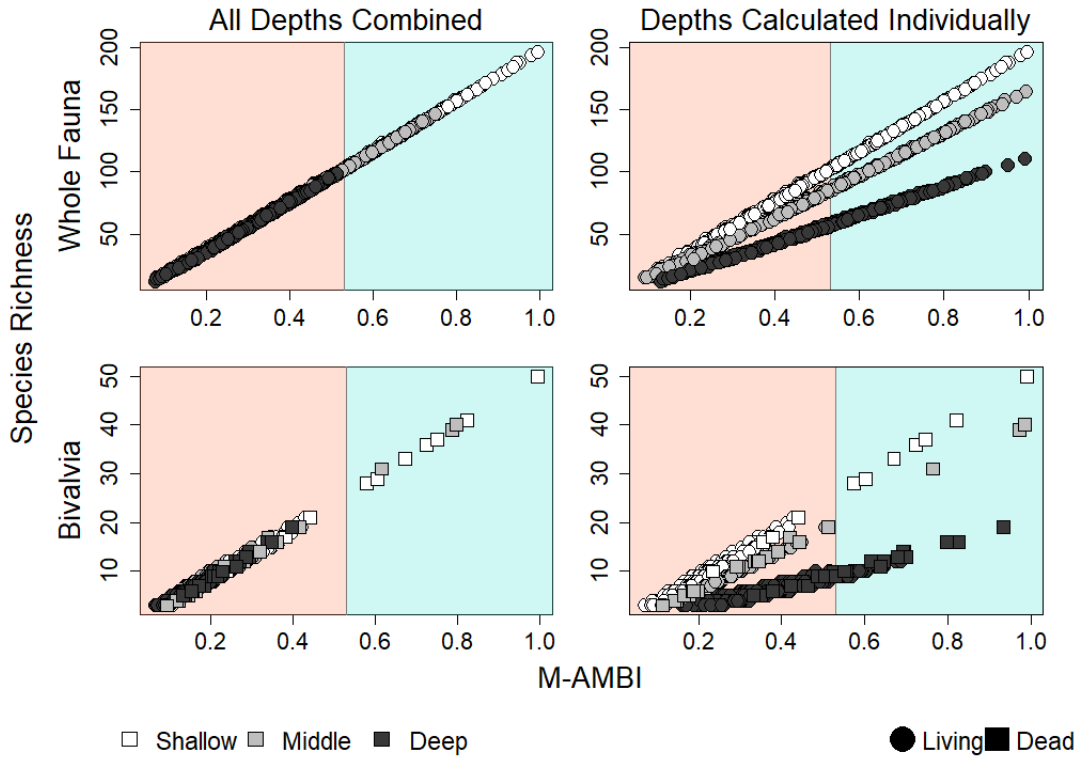

Figure F.8: Multivariate AMBI (M-AMBI) scores versus species richness for the whole fauna (top row) and bivalves (bottom row). Note that M-AMBI ranges from values of zero to one, with higher scores indicating better ecological quality. A threshold value of 0.53 divides the index's 'Good' and 'Moderate' EQS, and was thus used as the threshold for my binary EQS scheme (see Table 3.1). M-AMBI scores were calculated in two ways: first, all samples from the three isobaths were assessed together, falling along a single reference gradient (left column), and second, samples were calculated separately for each isobath, basing scores from each isobath on their own reference gradient (right column). Although the calculation of M-AMBI considers species richness, Shannon diversity, and the AMBI score of each sample, extremely strong positive correlation between M-AMBI score and species richness indicated that this metric alone contributes most to M-AMBI (and may thus reliably approximate it) on the Palos Verdes shelf. However, the decline in species richness with increased water depth consequentially reduces the EQS of deep-water samples if all samples are calculated together versus separated by water depth. Further, the higher richness (and evenness) encountered in bivalve death assemblages (square points in bottom row) versus annually-sampled living assemblages (circular points), an expected property of time averaging, may artificially reduce the apparent ecological quality of live-collected samples. Thus, when using M-AMBI to estimate EQS, it is essential in any system to consider both (1) natural environmental gradients that might be conflated with known and suspected anthropogenic disturbance, and (2) the predicted effects of time averaging if data from death assemblages are being integrated with singly-sampled (non-averaged) data from living assemblages.

APPENDIX G

SIMULATION FRAMEWORK AND PILOT RESULTS FOR BUILDING A DEAD-SHELL AGE MODEL FOR PUGET SOUND

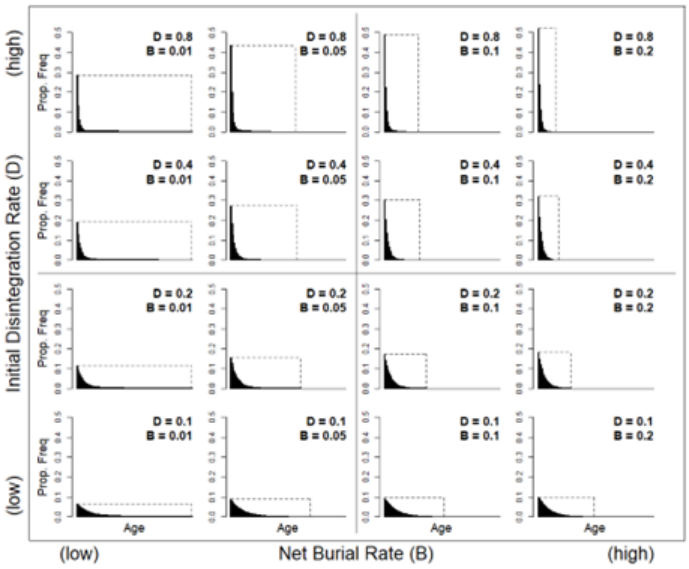

Figure G.1: Predicted effects on the shape of a shell age frequency distribution (AFD) in the surface mixed layer (SML) from the interaction of shell disintegration (aka loss) and net sedimentation (aka burial). In each graph, the y-axis is the proportional frequency of shells and the x-axis is the elapsed time since death, based on 1,000 time steps (not labeled). Dashed vertical lines indicate the age of the oldest surviving shell in the simulation and dashed horizontal lines indicate the proportional abundance of the youngest (most abundant) cohort in the assemblage, which is usually the first bin. Together, the dashed lines create a box that summarizes the shape of the AFD. Simulations are based on a 2-phase exponential model of shell loss (Tomašových et al., 2014) – that is, disintegration is initially fast and then, after some elapsed time to sequestration, slows sharply – and assumed a constant rate of shell production (i.e., same initial number of shells input to the SML per age cohort). Low sedimentation rates or other variables (such as bioturbation, i.e., vertical mixing) that reduce net burial promotes the retention of older shells in the assemblage. As net burial increases (towards the right), older shells are increasingly likely to be lost from the system via burial below the SML, reducing the length of time-averaging in the SML. The intensity of bioturbation could be added as a third axis: by advecting buried shells back up toward the sediment-water interface, they would be subjected to additional opportunities for shell disintegration from acidic porewaters and physical reworking along with newly introduced young shells; advection would also move very young shells down to the base of the SML, temporarily sequestering them or at the least admixing them with older shells there. Puget Sound, hypothesized to experience high disintegration rates from cold water and high net burial from high sedimentation ($\sim$2 cm/yr), is most likely an example of a system that would plot in the upper left of this phase space, with a steep initial slope and a relatively short right-skew.

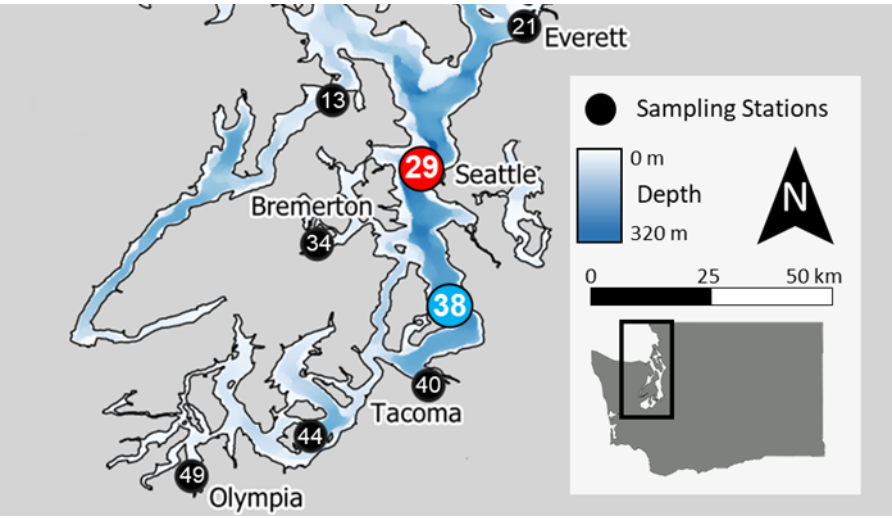

Figure G.2: Two sampling stations of shells for age-dating (rep point labeled 29 and blue point labeled 38) located in the Central Basin of Puget Sound. These stations are both characterized by similarly silty sediments in relatively deep water (200 m), and have had relatively large and steady populations of the bivalve *Macoma carlottensis* over the past 30 years (Figure G.3).

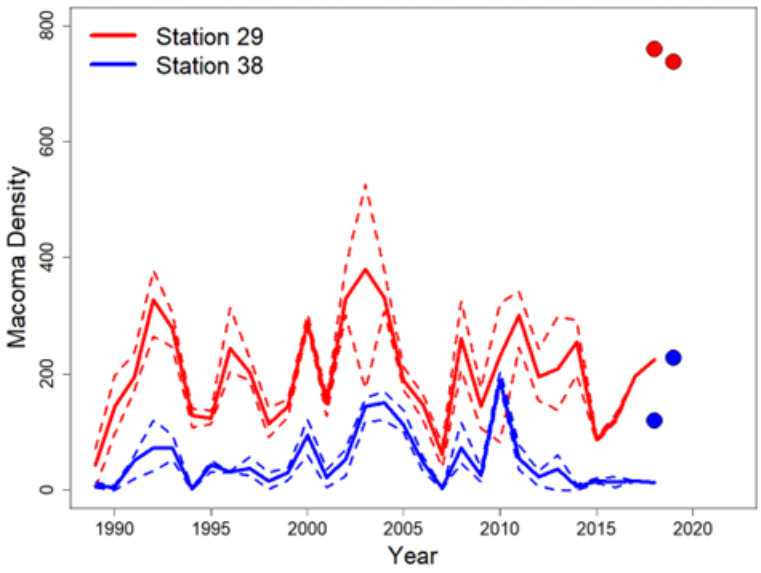

Figure G.3: Time series of population densities of *M. carlottensis* at the monitoring stations 29 and 38. Solid lines are the mean density (number of live-collected individuals per 0.1 m^2) among all available replicate samples for a given year (number of replicates varies from 1 to 5 for a given year). The dashed lines are the minimum and maximum observed densities among available replicates. Filled circles represent the density of individuals per van Veen grab from death assemblages sampled in 2018 and 2019. Living (and dead) densities were consistently higher at Station 29 than at Station 38.

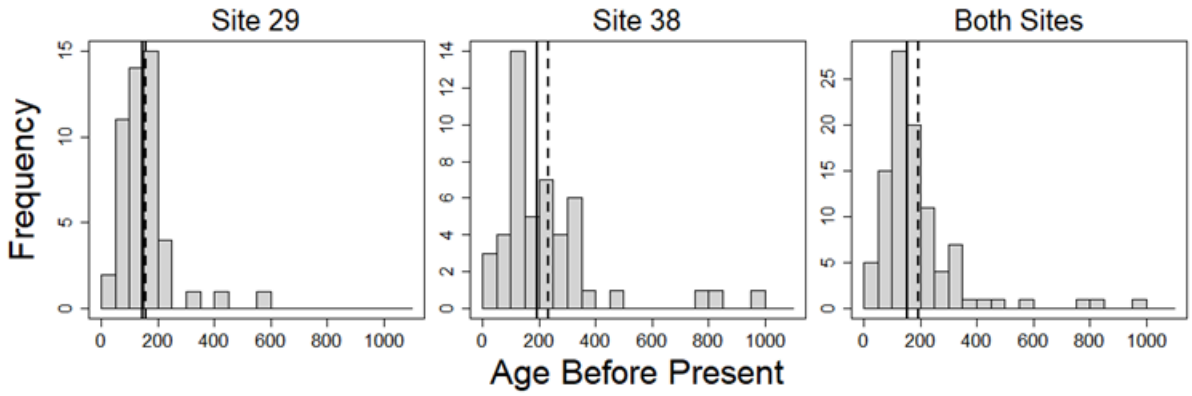

Figure G.4: Age-frequency distributions of the raw, non-corrected radiocarbon data (n = 50 at each site). Bars are divided into 50-year intervals. Solid vertical lines represent the median age and the dashed line represents the 75% interquartile range.

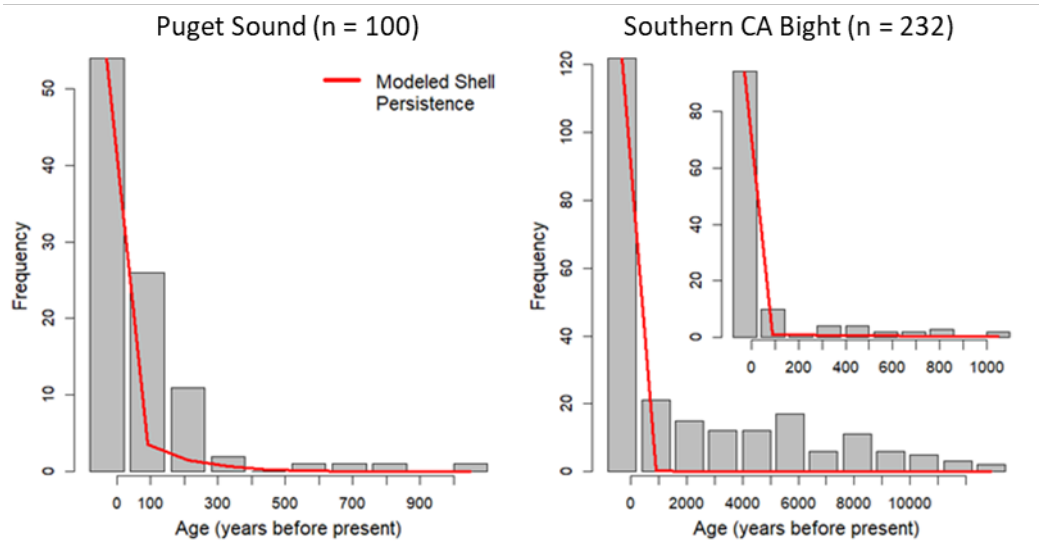

Figure G.5: Pilot results of AFDs for death assemblages from soft-sediment habitats from (left) Puget Sound, WA, with a maximum shell age of 1,100 years before present (non-corrected), and (right) from the Southern California Bight, with a maximum shell age of 11,000 years before present (data from Tomašových et al., 2014). Inset plot: distribution of 124 California shells corresponding to the same total age range (first 1,100 years) that characterizes the Puget Sound AFD. Red lines are the fitted model outputs for each dataset using the two-phase loss model of Tomašových et al. (2014), indicating a rapid initial rate of loss of very young shells (steep slope); shells surviving the first 100 years of residence in the SML then exhibit a much slower (two orders-of-magnitude lower) loss rate, having been effectively sequestered (diagenetic stabilization is suspected for California shells).

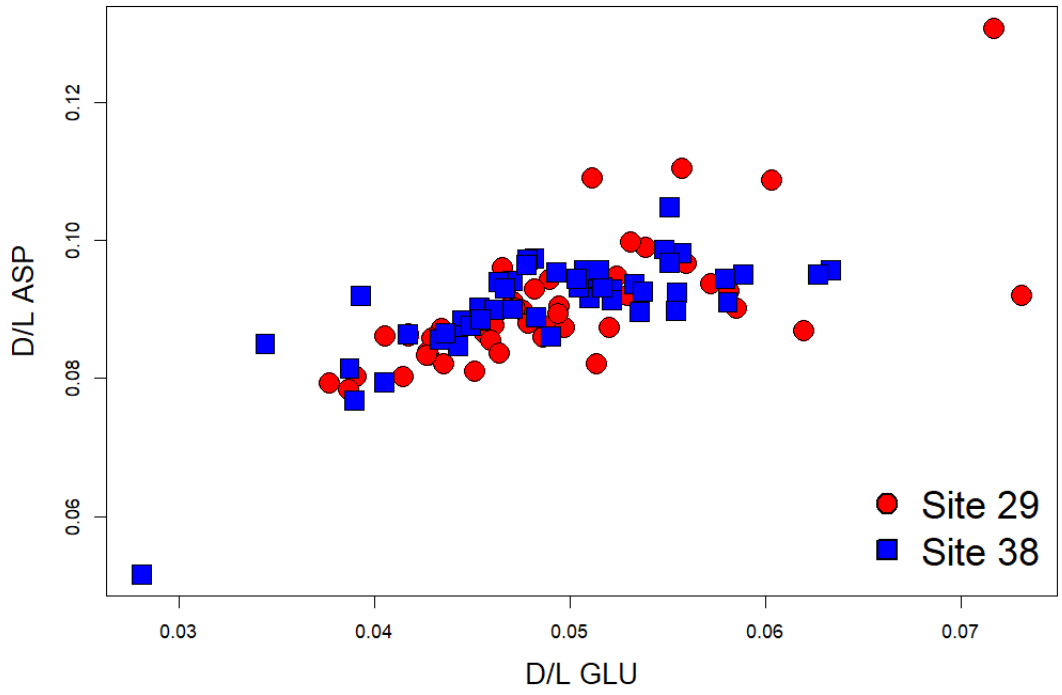

Figure G.6: The relationship between the D/L ratios of aspartic acid (ASP) and glutamic acid (GLU) measured within shells. Although the relationship demonstrates a positive correlation, anomalous endmembers appear to drive the trend.

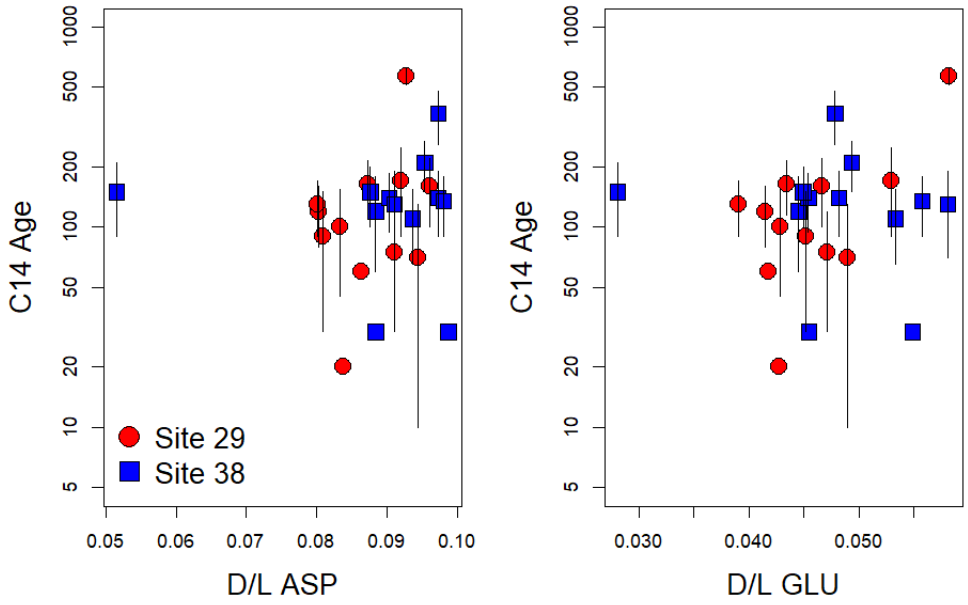

Figure G.7: The relationship between preliminary (non-corrected) radiocarbon ages of shells and the D/L ratios for aspartic acid (left) and glutamic acid (right) of shells from both sites. Black lines are error bars for radiocarbon ages (note logged vertical axis).

REFERENCES

Agard J., Kishore R., and Bayne B. Perna viridis (linnaeus, 1758): first record of the indo-pacific green mussel (mollusca: Bivalvia) in the caribbean. *Caribbean Marine Studies*, 3: 59–60, 1992.

Albano P.G. Comparison between death and living land mollusk assemblages in six forested habitats in northern italy. *Palaios*, 29(7):338–347, 2014. doi:10.2110/palo.2014.020.

Albano P.G., Gallmetzer I., Haselmair A., Tomašových A., Stachowitsch M., and Zuschin M. Historical ecology of a biological invasion: The interplay of eutrophication and pollution determines time lags in establishment and detection. *Biological Invasions*, 20:1417–1430, 2018. doi:10.1007/s10530-017-1634-7.

Albano P.G., Hua Q., Kaufman D.S., Tomašových A., Zuschin M., and Agiadi K. Radiocarbon dating supports bivalve-fish age coupling along a bathymetric gradient in high-resolution paleoenvironmental studies. *Geology*, 48(6):589–593, 2020. doi:10.1130/G47210.1.

Albano P.G., Steger J., Bošnjak M., Dunne B., Guifarro Z., Turapova E., Hua Q., Kaufman D.S., Rilov G., and Zuschin M. Native biodiversity collapse in the eastern mediterranean. *Proceedings of the Royal Society B*, 288:20202469, 2021. doi:10.1098/rspb.2020.2469.

Albano P.G., Tomašových A., Stachowitsch M., and Zuschin M. Taxonomic sufficiency in a live-dead agreement study in a tropical setting. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 499:341–348, 2016. doi:10.1016/j.palaeo.2016.02.031.

Alleng G.P. The fauna of the port royal mangal, kingston, jamaica. *Studies on the Natural History of the Caribbean Region*, 73(1):23–42, 1997.

Alonso-Hernández C.M., Gómez-Batista M., Cattini C., Villeneuve J-P., and J. Oh. Organochlorine pesticides in green mussel, perna viridis, from the cienfuegos bay, cuba. *Bulletin of Environmental Contamination and Toxicology*, 89(5):995–999, 2012. doi:10.1007/s00128-012-0835-0.

Anderson M.J. A new method for non-parametric multivariate analysis of variance. *Bulletin of Environmental Contamination and Toxicology*, 26(1):32–46, 2001. doi:10.1111/j.1442-9993.2001.01070.pp.x.

Anderson M.J., Ellingsen K.E., and McArdle B.H. Multivariate dispersion as a measure of beta diversity. *Ecology Letters*, 9(6):683–693, 2006. doi:10.1111/j.1461-0248.2006.00926.x.

Aslan H. and Ovalis P. Spatial variation ecological quality assessment of live and dead molluscan assemblages in the aegean sea. *Ocean Science Journal*, 58(3):20, 2023. doi:10.1007/s12601-023-00114-1.

Atmore L.M., Aiken M., and Furni F. Shifting ecological baselines to thresholds: reframing exploitation in the marine environment. *Frontiers in Marine Science*, 8:742188, 2021. doi:10.3389/fmars.2021.742188.

Bacci T., Trabucco B., Marzialetti S., Marusso V., Lomiri S., Vani D., and Lamberti C.V. Taxonomic sufficiency in two case studies: where does it work better? *Marine Ecology*, 30(s1):13–19, 2009. doi:10.1111/j.1439-0485.2009.00324.x.

Baker P., Fajans J.S., Arnold W.S., Ingrao D.A., Marelli D.C., and Baker S.M. Range and dispersal of a tropical marine invader, the asian green mussel, perna viridis, in subtropical waters of the southeastern united states. *Journal of Shellfish Research*, 26:345–355, 2007. doi:10.2983/0730-8000(2007)26[345:RADOAT]2.0.CO;2.

Behrensmeyer A.K. and Kidwell S.M. Taphonomy's contributions to paleobiology. *Paleobiology*, 11(1):105–119, 1985. doi:10.1017/S009483730001143X.

Bertasi F., Colangelo M.A., Colosio F., Gregorio G., Abbiati M., and Ceccherelli V.U. Comparing efficacy of different taxonomic resolutions and surrogates in detecting changes in soft bottom assemblages due to coastal defence structures. *Marine Pollution Bulletin*, 58 (5):686–694, 2009. doi:10.1016/j.marpolbul.2009.01.003.

Berthelsen A, Atalah J., Clark D., Goodwin E., Patterson M., and Sinner J. Relationships between biotic indices, multiple stressors and natural variability in new zealand estuaries. *Ecological Indicators*, 85:634–643, 2018. doi:10.1016/j.ecolind.2017.10.060.

Betancourt J.L., Long A., Donahue D.J., Jull A.J.T., and Zabel T.H. Pre-columbian age for north american corispermum l. (chenopodiaceae) confirmed by accelerator radiocarbon dating. *Nature*, 311:353–655, 1984. doi:10.1038/311653a0.

Bevilacqua S., Fraschetti S., Musco L., and Terlizzi A. Taxonomic sufficiency in the detection of natural and human-induced changes in marine assemblages: a comparison of habitats and taxonomic groups. *Marine Pollution Bulletin*, 58(12):1850–1859, 2009. doi:10.1016/j.marpolbul.2009.07.018.

Terlizzi S., Bevilacqua A., Claudet J., Fraschetti S., and Boero F. Taxonomic relatedness does not matter for species surrogacy in the assessment of community responses to environmental drivers. *Journal of Applied Ecology*, 49(2):357–366, 2012. doi:10.1111/j.1365-2664.2011.02096.x.

Bhusal D.R., Kallimanis A.S., Tsiafouli M.A., and Sgardelis S.P. Higher taxa vs. functional guilds vs. trophic groups as indicators of soil nematode diversity and community structure. *Ecological Indicators*, 41:25–29, 2014. doi:10.1016/j.ecolind.2014.01.019.

Bizjack J.L., Kidwell S.M., Velarde R.D., Leonard-Pingel J., and Tomašových A. Detecting, sourcing, and age-dating dredged sediments on the open shelf, southern california, using dead mollusk shells. *Marine Pollution Bulletin*, 114(1):448–465, 2017. doi:10.1016/j.marpolbul.2016.10.010.

Booth D.B. Glaciofluvial infilling and scour of the puget lowland, washington, during ice-sheet glaciation. *Geology*, 22(8):695–698, 1994.

Borja Á., Chust G., Rodríguez J.G., Bald J., Belzunce-Segarra M.J., Franco J., Garmendia J.M., Larreta J., Menchaca I., Muxika I., Solaun O., Revilla M., Uriarte A., Valencia V., and Zorita I. The past is the future of the present: learning from long-time series of marine monitoring. *Science of the Total Environment*, 566-567:698–711, 2016. doi:10.1016/j.scitotenv.2016.05.111.

Borja A., Chust G., and Muxika I. Chapter three – forever young: the successful story of a marine biotic index. *Advances in Marine Biology*, 82:93–127, 2019a. doi:10.1016/bs.amb.2019.05.001.

Borja Á., Dauer D.M., and Grémare A. The importance of setting targets and reference conditions in assessing marine ecosystem quality. *Ecological Indicators*, 12(1):1–7, 2012. doi:10.1016/j.ecolind.2011.06.018.

Borja A., Franco J., and Pérez V. The importance of setting targets and reference conditions in assessing marine ecosystem quality. *Marine Pollution Bulletin*, 40(12):1100–1114, 2000. doi:10.1016/S0025-326X(00)00061-8.

Borja A. and Muxika I. Guidelines for the use of ambi (atzi's marine biotic index) in the assessment of the benthic ecological quality. *Marine Pollution Bulletin*, 50(7):787–789, 2005. doi:10.1016/j.marpolbul.2005.04.040.

Borja Á., Muxika I., and Franco J. Long-term recovery of soft-bottom benthos following urban and industrial sewage treatment in the nervión estuary (southern bay of biscay). *Marine Ecology Progress Series*, 313:43–55, 2006. doi:10.3354/meps313043.

Boudreaux M.L. and Walters L.J. Mytella charruana (bivalvia: Mytillidae): a new, invasive bivalve in mosquito lagoon, florida. *The Nautilus*, 120(1):34–36, 2006.

Brandenberger J.M., Louchouarn P., and Crecelius E.A. Natural and post-urbanization signatures of hypoxia in two basins of puget sound: historical reconstruction of redox sensitive metals and organic matter inputs. *Aquatic Geochemistry*, 17:645–670, 2011. doi:10.1007/s10498-011-9129-0.

Bray J.R. and Curtis J.T. An ordination of the upland forest communities of southern wisconsin. *Ecological Monographs*, 27(4):325–349, 1957. doi:10.2307/1942268.

Brind'Amour A., Laffargue P., Morin J., Vaz S., Foveau A., and Le Bris H. Morphospecies and taxonomic sufficiency of benthic megafauna in scientific bottom trawl surveys. *Continental Shelf Research*, 72:1–9, 2014a. doi:10.1016/j.csr.2013.10.015.

Bromley R.G. and Heinberg C. Attachment strategies of organisms on hard substrates: a palaeontological view. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 223(2-4):429–453, 2006. doi:10.1016/j.palaeo.2005.07.007.

Buddo D.S.A. *The biology and ecology of the invasive Indo-Pacific green mussel, Perna viridis in Kingston Harbour, Jamaica.* Phd thesis, University of the West Indies, Mona, Jamaica, 2008.

Buddo D.S.A., Steele R.D., and D'Oyen E.R. Distribution of the invasive indo-pacific green mussel, perna viridis, in kingston harbour, jamaica. *Bulletin of Marine Science*, 73(2): 433–441, 2003.

Buddo D.S.A., Steele R.D., and Webber M.K. Public health risks posed by the invasive indo-pacific green mussel, perna viridis (linnaeus, 1758) in kingston harbour, jamaica. *BioInvasions Records*, 1(3):171–178, 2012. doi:10.3391/bir.2012.1.3.03.

Buss D.F. and Vitorino A.S. Rapid bioassessment protocols using benthic macroinvertebrates in brazil: evaluation of taxonomic sufficiency. *Journal of the North American Benthological Society*, 29(2):562–571, 2012. doi:10.1899/09-095.1.

Cai W., Meng W., Zhu Y., Zhou J., and Liu L. Assessing benthic ecological status in stressed liaodong bay (china) with ambi and m-ambi. *Chinese Journal of Oceanology and Limnology*, 31:482–492, 2013. doi:10.1007/s00343-013-2177-0.

Carneiro F.M., Bini L.M., and Rodrigues L.C. Influence of taxonomic and numerical resolution on the analysis of temporal changes in phytoplankton communities. *Ecological Indicators*, 10(2):249–255, 2010. doi:10.1016/j.ecolind.2009.05.004.

Carter J.G. *Skeletal Biomineralization: Patterns, Processes, and Evolutionary Trends Volume 1.* Van Nostrand Reinhold, 1990.

Casebolt S. and Kowalewski M. Mollusk shell assemblages as archives of spatial structuring of benthic communities around subtropical islands. *Estuarine, Coastal and Shelf Science*, 215:132–143, 2018. doi:10.1016/j.ecss.2018.09.023.

Caswell B.A., Paine M., and Frid C.L.J. Seafloor ecological functioning over two decades of organic enrichment. *Marine Pollution Bulletin*, 136:212–229, 2018. doi:10.1016/j.marpolbul.2018.08.041.

Chiba T. and Sato S.I. Invasion of laguncula pulchella (gastropoda: Naticidae) and predator-prey interactions with bivalves on the tona coast, miyagi prefecture, northern japan. *Biological Invasions*, 15:587–598, 2013. doi:10.1007/s10530-012-0310-1.

City of San Diego. *Biennial Receiving Waters Monitoring and Assessment Report for the Point Loma and South Bay Ocean Outfalls 2020-2021.* City of San Diego Ocean Monitoring Program, San Diego, CA, 2022.

Clare D.S., Robinson L.A., and Frid C.L.J. Community variability and ecological functioning: 40 years of change in the north sea benthos. *Marine Environmental Research*, 107:24–34, 2015. doi:10.1016/j.marenvres.2015.03.012.

Clarke K.R. Non-parametric multivariate analyses of changes in community structure. *Australian Journal of Ecology*, 18(1):117–143, 1993. doi:10.1111/j.1442-9993.1993.tb00438.x.

Clarke K.R. and Warwick R.M. *Changes in marine communities: an approach to statistical analysis and interpretation, 2nd Edn.* PRIMER-E, Plymouth, 2001.

Close R.A., Evers S.W., Alroy J., and Butler R.J. How should we estimate diversity in the fossil record? testing richness estimators using sampling-standardised discovery curves. *Methods in Ecology and Evolution*, 9(6):1386–1400, 2018. doi:10.1111/2041-210X.12987.

Coan E.V., Scott P.V., and Bernard F.R. *Bivalve seashells of western North America.* Santa Barbara Museum of Natural History Monographs 2, Santa Barbara, CA, 2000.

Coffey E.E.D., Froyd C.A., and Willis K.J. When is an invasive not an invasive? macrofossil evidence of doubtful native plant species in the galápagos islands. *Ecology*, 92(4):805–812, 2011. doi:10.3389/fmars.2016.00169.

Collins K.S., Edie S.M., Gao T., Bieler R., and Jablonski D. Spatial filters of function and phylogeny determine morphological disparity with latitud. *PLOS ONE*, 14:e0221490, 2019. doi:10.1371/journal.pone.0221490.

De Oliveira S.S. Jr., Ortega J.C.G., dos Santos Ribas L.G., Lopes V.G., and Bini L.M. Higher taxa are sufficient to represent biodiversity patterns. *Ecological Indicators*, 111: 105994, 2020. doi:10.1016/j.ecolind.2019.105994.

De Messano L.V.R., Gonçalves J.E.A., Messano H.F., Campos S.H.C., and Coutinho R. First report of the asian green mussel perna viridis (linnaeus, 1758) in rio de janeiro, brazil: a new record for the southern atlantic ocean. *BioInvasions Records*, 8:653–660, 2019. doi:10.3391/bir.2019.8.3.22.

Dethier M.N. and Schoch G.C. Taxonomic sufficiency in distinguishing natural spatial patterns on an estuarine shoreline. *Marine Ecology Progress Series*, 306:41–49, 2006. doi:10.3354/meps306041.

Di Minin E. and Moilanen A. Improving the surrogacy effectiveness of charismatic megafauna with well-surveyed taxonomic groups and habitat types. *Journal of Applied Ecology*, 51 (2):281–288, 2014. doi:10.1111/1365-2664.12203.

Dias P.J., Gilg M.R., Lukehurst S.S., Kennington W.J., Huhn M., Madduppa H.H., McKirdy S.J., de Lestang P., Teo S.L.M., Lee S.S.C., and McDonald J.I. Genetic diversity of a hitchhiker and prized food source in the anthropocene: the asian green mussel perna viridis (mollusca, mytilidae). *Biological Invasions*, 20:1749–1770, 2018. doi:10.1007/s10530-018-1659-6.

Dietl G.P., Durham S.R., Smith J.A., and Tweitmann A. Mollusk assemblages as records of past and present ecological status. *Frontiers in Marine Science*, 3:169, 2016. doi:10.3389/fmars.2016.00169.

Dietl G.P. and Flessa K.W. Conservation paleobiology: putting the dead to work. *Trends in Ecology and Evolution*, 26(1):30–37, 2011. doi:10.1016/j.tree.2010.09.010.

Dietl G.P., Kidwell S.M., Brenner M., Burney D.A., Flessa K.W., Jackson S.T., and Koch P.L. Conservation paleobiology: leveraging knowledge of the past to inform conservation and restoration. *Annual Review of Earth and Planetary Sciences*, 43:79–103, 2015. doi:10.1146/annurev-earth-040610-133349.

Dillon E.M., Pier J.Q., J.A. Smith, N.B. Raja, Dimitrijević D., Austin E.L., Cybulski J.D., De Entrambasaguas J., Durham S.R., Grether C.M., Haldar H.S., Kocákova K., Lin C-H., Mazzini I., Mychajliw A.M., Ollendorf A.L., Pimiento C., Fernández O.R.R., Smith I.E., and Dietl G.P. What is conservation paleobiology? tracking 20years of research and development. *Frontiers in Ecology and Evolution*, 10:1031483, 2022. doi:10.3389/fevo.2022.1031483.

Dimitriou P.D., Apostolaki E.T., Papageorgiou N., Reizopoulou S., Simboura N., Arvantidis C., and Karakassis I. Meta-analysis of a large data set with water framework directive indicators and calibration of a benthic quality index at the family level. *Ecological Indicators*, 20:101–107, 2012. doi:10.1016/j.ecolind.2012.02.008.

Domínguez-Castanedo N., Rojas-López R., Solís-Weiss V., Hernández-Alcántara P., and Granados-Barba A. The use of higher taxa to assess the benthic conditions in the southern gulf of mexico. *Marine Ecology*, 28(s1):161–168, 2007. doi:10.1111/j.1439-0485.2007.00178.x.

Dutch M., Partridge V., Weakland S., Burgess D., and Eagleston A. *Quality assurance monitoring plan: the Puget Sound sediment monitoring program.* Washington State Department of Ecology Publication, Lacey, WA, 2018.

Eganhouse R.P. and Pontolillo J. Depositional history of organic contaminants on the palos verdes shelf, california. *Marine Chemistry*, 70(4):317–338, 2000. doi:10.1016/S0304-4203(00)00033-5.

Eleftheriou A. and Moore D.C. Macrofauna techniques. In Eleftheriou A., editor, *Methods for the study of marine benthos*, pages 175–251. John Wiley and Sons, Chichester, 2013.

Elliott T., Persad G., and Webber M. Variation in the colonization of artificial substrates by mangrove root fouling species of the port royal mangrove lagoons in the eutrophic kingston harbour, jamaica. *Journal of Water Resource and Protection*, 4(6):377–387, 2012. doi:10.4236/jwarp.2012.46043.

Ellis D. Taxonomic sufficiency in pollution assessment. *Marine Pollution Bulletin*, 16(12): 459, 1985. doi:10.1016/0025-326X(85)90362-5.

Fabrikant R. The effect of sewage effluent on the population density and size of the clam parvilucina tenuisculpta. *Marine Pollution Bulletin*, 15(7):249–253, 1984. doi:10.1016/0025-326X(84)90363-1.

Ferraro S.P. and Cole F.A. Taxonomic level sufficient for assessing a moderate impact on macrobenthic communities in puget sounds washington, usa. *Canadian Journal of Fisheries and Aquatic Sciences*, 49(6):1184–1188, 1992. doi:10.1139/f92-133.

Fesser K.M. and Miller A.I. Temporal dynamics of shallow seagrass-associated molluscan assemblages in st. croix, u.s. virgin islands: toward the calibration of taphonomic inertia. *Palaios*, 29(5):218–230, 2014. doi:10.2110/palo.2013.103.

Flenley J.R., King A.S.M., Jackson J., Chew C., Teller J.T., and Prentice M.E. The late quaternary vegetational and climatic history of easter island. *Journal of Quaternary Science*, 6(2):85–115, 1991. doi:10.1002/jqs.3390060202.

Fontaine A., Devillers R., Peres-Neto P.R., and Johnson L.E. Delineating marine ecological units: a novel approach for deciding which taxonomic group to use and which taxonomic resolution to choose. *Diversity and Distributions*, 21(10):1167–1180, 2015. doi:10.1111/ddi.12361.

Forcino F.L., Stafford E.S., and Leighton L.R. Perception of paleocommunities at different taxonomic levels: how low must you go? *Palaeogeography, Palaeoclimatology, Palaeoecology*, 365-366:48–56, 2012. doi:10.1016/j.palaeo.2012.09.011.

Francis P.A., Maxam S.A., and Webber M.K. Rapid reassessment of the eutrophication status of kingston harbour, jamaica using the zooplankton community. *Revista de Biología Tropical*, 62:223–239, 2014. doi:10.15517/rbt.v62i0.15918.

Fürsich F.T. and Aberhan M. Significance of time-averaging for palaeocommunity analysis. *Lethaia*, 23:143–152, 1990. doi:10.1111/j.1502-3931.1990.tb01355.x.

Giangrande A., Licciano M., and Musco L. Polychaetes as environmental indicators revisited. *Marine Pollution Bulletin*, 50(11):1153–1162, 2005. doi:10.1016/j.marpolbul.2005.08.003.

Gilad E., Kidwell S.M., Benayuha Y., and Edelman-Furstenberg Y. Unrecognized loss of seagrass communities based on molluscan death assemblages: historic baseline shift in tropical gulf of aqaba, red sea. *Marine Ecology Progress Series*, 589:73–83, 2018. doi:10.3354/meps12492.

Gilg M.R., Johnson E.G., Gobin J., Bright B.M., and Ortolaza A.I. Population genetics of introduced and native populations of the green mussel, perna viridis: determining patterns of introduction. *Biological Invasions*, 15(2):73–83, 2014. doi:10.1007/s10530-012-0301-2.

Gilinski N.L. and Bennington J.B. Estimating numbers of whole individuals from collections of body parts: a taphonomic limitation of the paleontological record. *Paleobiology*, 20(2): 245–258, 1994. doi:10.1017/S0094837300012719.

Gillett D.J., Lovell L.L., and Schiff K.C. *Southern California Bight 2013 Regional Monitoring Program: Volume VI Benthic Infauna.* Southern California Coastal Water Research Project Technical Report 971, Costa Mesa, CA, 2017.

Gillett D.J., Weisberg S.B., Grayson T., Hamilton A., Hansen V., Leppo E.W., M.C. Pelletier, Borja A., Cadien D., Dauer D., Diaz R., Dutch M., Hyland J.L., Kellogg M., Larsen P.F., Levinton J.S., Llansó R., Lovell L.L., Montagna P.A., Pasko D., Philips C.A., Rakocinski C., Ranasinghe J.A., Sanger D.M., Teixeira H., Van Dolah R.F., Velarde R.G., and Welch K.I. Effects of ecological group classification schemes on performance of the ambi benthic index in us coastal waters. *Ecological Indicators*, 50:99–107, 2015. doi:10.1016/j.ecolind.2017.08.067.

Gladstone W., Murray B.R., and Hutchings P. Promising yet variable performance of cross-taxon biodiversity surrogates: a test in two marine habitats at multiple times. *Biodiversity and Conservation*, 29:3067–3089, 2020. doi:10.1007/s10531-020-02015-4.

Gobin J., Agard J., Madera J., and Mohammed A. The asian green mussel perna viridis (linnaeus 1758): 20 years after its introduction in trinidad and tobago. *Open Journal of Marine Science*, 3(2):62–65, 2013. doi:10.4236/ojms.2013.32007.

Goodbody I. Kingston harbour, jamaica – an overview. *Bulletin of Marine Science*, 73(2): 249–255, 2003. doi:10.4236/ojms.2013.32007.

Gogina M., Darr A., and Zettler M.L. Approach to assess consequences of hypoxia disturbance events for benthic ecosystem functioning. *Journal of Marine Systems*, 129:203–213, 2014. doi:10.1016/j.jmarsys.2013.06.001.

Gracia A.C. and Rangel-Buitrago N. The invasive species perna viridis (linnaeus, 1758 - bivalvia: Mytilidae) on artificial substrates: a baseline assessment for the colombian caribbean sea. *Marine Pollution Bulletin*, 152:110926, 2020. doi:10.1016/j.marpolbul.2020.110926.

Gray J.S. and Christie H. Predicting long-term changes in marine benthic communities. *Marine Ecology Progress Series*, 13(1):87–94, 1983.

Gray J.S. and Elliott M. *Ecology of marine sediments: from science to management, 2nd edn.* Oxford University Press, New York, NY, 2009.

Green S.J. and Grosholz E.D. Functional eradication as a framework for invasive species control. *Frontiers in Ecology and the Environment*, 19(2):98–107, 2020. doi:10.1002/fee.2277.

Gusmao J.B., Brauko K.M., Eriksson B.K., and Lana P.C. Functional diversity of macrobenthic assemblages decreases in response to sewage discharges. *Ecological Indicators*, 66:65–75, 2016. doi:10.1016/j.ecolind.2016.01.003.

Harnik P.G. Unveiling rare diversity by integrating museum, literature, and field data. *Paleobiology*, 35(2):190–208, 2009. doi:10.1666/07062.1.

Haselmair A., Gallmetzer I., Tomašových A., Wieser A.M., Übelhör A., and Zuschin M. Basin-wide infaunalisation of benthic soft-bottom communities driven by anthropogenic habitat degradation in the northern adriatic sea. *Marine Ecology Progress Series*, 671: 45–65, 2021a. doi:10.3354/meps13759.

Heino J. Are indicator groups and cross-taxon congruence useful for predicting biodiversity in aquatic ecosystems? *Ecological Indicators*, 10(2):112–117, 2010. doi:10.1016/j.ecolind.2009.04.013.

Heino J. Taxonomic surrogacy, numerical resolution and responses of stream macroinvertebrate communities to ecological gradients: are the inferences transferable among regions? *Ecological Indicators*, 36:186–194, 2014. doi:10.1016/j.ecolind.2013.07.022.

Heino J. and Soininen J. Are higher taxa adequate surrogates for species-level assemblage patterns and species richness in stream organisms? *Biological Conservation*, 137(1):78–89, 2007. doi:10.1016/j.biocon.2007.01.017.

Hendry M.D. Historical evidence of shoreline evolution for the palisadoes, kingston. *Journal of the Geological Society of Jamaica*, 18:39–42, 1978.

Huang Y.C., Li Z.K., Chen W.L., Chan C.C., Hsu H.Y., Lin Y.T., Huang Y.S., and Han Y.S. First record of the invasive biofouling mussel mytella strigata (hanley, 1843) (bivalvia: Mytilidae) from clam ponds in taiwan. *BioInvasions Records*, 10(2):304–312, 2021. doi:10.3391/bir.2021.10.2.08.

Hulme P.E. Trade, transport and trouble: managing invasive species pathways in an era of globalization. *Journal of Applied Ecology*, 46(1):10–18, 2009. doi:10.1111/j.1365-2664.2008.01600.x.

Hurlbert S.H. The nonconcept of species diversity: a critique and alternative parameters. *Ecology*, 52(4):577–586, 1971. doi:10.2307/1934145.

Hyman A.C., Frazer T.K., Jacoby C.A., Frost J.R., and Kowalewski M. Long-term persistence of structured habitats: seagrass meadows as enduring hotspots of biodiversity and faunal stability. *Proceedings of the Royal Society B*, 286:20161861, 2019. doi:10.1098/rspb.2019.1861.

Ingrao D.A., Mikkelsen P.M., and Hicks D.W. Another introduced marine mollusk in the gulf of mexico: the indo-pacific green mussel, perna viridis, in tampa bay, florida. *Journal of Shellfish Research*, 20(1):13–19, 2001.

Islam M.S. and Tanaka M. Impacts of pollution on coastal and marine ecosystems including coastal and marine fisheries and approach for management: a review and synthesis. *Marine Pollution Bulletin*, 48(7-8):624–649, 2004. doi:10.1016/j.marpolbul.2003.12.004.

Jayachandran P.R., Aneesh B.P., Oliver P.G., Philomina J., Jima M., Harikrishnan K., and Nandan S.B. First record of an alien invasive biofouling mussel mytella strigata (hanley, 1843) (mollusca: Mytilidae) from indian waters. *BioInvasions Records*, 8(4):13–19, 2019. doi:10.3391/bir.2019.8.4.11.

Jones F.C. Taxonomic sufficiency: the influence of taxonomic resolution on freshwater bioassessments using benthic macroinvertebrates. *Environmental Reviews*, 16:45–69, 2008. doi:10.1139/A07-010.

Khangaonkar T., Yang Z., Kim T., and Roberts M. Tidally averaged circulation in puget sound sub-basins: comparison of historical data, analytical model, and numerical model. *Estuarine, Coastal and Shelf Science*, 93(4):305–319, 2011. doi:10.1016/j.ecss.2011.04.016.

Kaufman D.S. and Manley W.F. A new procedure for determining dl amino acid ratios in fossils using reverse phase liquid chromatography. *Quaternary Science Reviews*, 17(11): 987–1000, 1998. doi:10.1016/S0277-3791(97)00086-3.

Kidwell S.M. Preservation of species abundance in marine death assemblages. *Science*, 294 (5544):1091–1094, 2001. doi:10.1126/science.1064539.

Kidwell S.M. Mesh-size effects on the ecological fidelity of death assemblages: a meta-analysis of molluscan live-dead studies. *Geobios*, 35(7):107–119, 2002. doi:10.1016/S0016-6995(02)00052-9.

Kidwell S.M. Discordance between living and death assemblages as evidence for anthropogenic ecological change. *Proceedings of the National Academy of Sciences*, 104(45): 17701–17706, 2007. doi:10.1073/pnas.0707194104.

Kidwell S.M. Ecological fidelity of open marine molluscan death assemblages: effects of post-mortem transportation, shelf health, and taphonomic inertia. *Lethaia*, 41(3):199–217, 2008. doi:10.1111/j.1502-3931.2007.00050.x.

Kidwell S.M. Evaluating human modification of shallow marine ecosystems: mismatch in composition of molluscan living and time-averaged death assemblages. In Dietl G.P. and Flessa K.W., editors, *Conservation Paleobiology: Using the Past to Manage for the Future*, pages 113–139. The Paleontological Society Papers 15, Portland, OR, 2009.

Kidwell S.M. Time-averaging and fidelity of modern death assemblages: building a taphonomic foundation for conservation paleobiology. *Palaeontology*, 56(3):487–522, 2013. doi:10.1111/pala.12042.

Kidwell S.M., Best M.M.R., and Kaufman D.S. Taphonomic trade-offs in tropical marine death assemblages: differential time averaging, shell loss, and probable bias in siliciclastic vs. carbonate facies. *Geology*, 33(9):729–732, 2005. doi:10.1130/G21607.1.

Kidwell S.M. and Tomasovych A. Implications of time-averaged death assemblages for ecology and conservation paleobiology. *Annual Review of Ecology, Evolution, and Systematics*, 44:539–563, 2013. doi:10.1146/annurev-ecolsys-110512-135838.

Klassen G. and Locke A. *A biological synopsis of the European green crab, Carcinus maenas*. Canadian Manuscript Report of Fisheries and Aquatic Sciences No. 2818, 2007.

Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. *International Joint Conference on Artificial Intelligence*, 14:1137–1145, 1995.

Kokesh B.S., Burgess D., Partridge V., Weakland S., and Kidwell S.M. Living and dead bivalves are congruent surrogates for whole benthic macroinvertebrate communities in puget sound. *Frontiers in Ecology and Evolution*, 10:980753, 2022a. doi:10.3389/fevo.2022.980753.

Kokesh B.S., Kidwell S.M., Tomašových A., and Walther S.M. Detecting strong spatial and temporal variation in macrobenthic composition on an urban shelf using taxonomic surrogates. *Marine Ecology Progress Series*, 682:13–30, 2022b. doi:10.3354/meps13932.

Kokesh B.S. and Stemann T.A. Dead men still tell tales: bivalve death assemblages record dynamics and consequences of recent biological invasions in kingston harbour, jamaica. *Geological Society, London, Special Publications*, 529(1):65–78, 2023. doi:10.1144/SP529-2022-28.

Korpanty C.A. and Kelley P.H. Molluscan live-dead agreement in anthropogenically stressed seagrass habitats: siliciclastic versus carbonate environments. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 410:113–125, 2014. doi:10.1016/j.palaeo.2014.05.014.

Kosnik M.A., Hua Q., Kaufman D.S., and Zawadzki A. sediment accumulation, stratigraphic order, and the extent of time-averaging in lagoonal sediments: a comparison of 210pb and 14c/amino acid racemization chronologies. *Coral Reefs*, 34:215–229, 2015. doi:10.1007/s00338-014-1234-2.

Kotta J., Lauringson V., and Kotta I. Response of zoobenthic communities to changing eutrophication in the northern baltic sea. *Hydrobiologia*, 580:97–108, 2007. doi:10.1007/s10750-006-0462-z.

Kowalewski M., Goodfriend G.A., and Flessa K.W. High-resolution estimates of temporal mixing within shell beds: the evils and virtues of time-averaging. *Paleobiology*, 24(3): 287–304, 1998. doi:10.1666/0094-8373(1998)024[0287:HEOTMW]2.3.CO;2.

Kowalewski M. and Hoffmeister A.P. Sieves and fossils: effects of mesh size on paleontological patterns. *Palaios*, 18(4-5):460–469, 2003. doi:10.1669/0883-1351(2003)018<0460:SAFEOM>2.0.CO;2.

Carroll M., Kowalewski M., Casazza L., Gupta N.S., Hannisdal B., Hendy A., Krause R.A. Jr., LaBarbera M., Lazo D.G., Messina C., Puchalski S., Rothfus T.A., Sälgeback J., Stempien J., Terry R.C., and Tomašových A. Quantitative fidelity of brachiopod-mollusk assemblages from modern subtidal environments of san juan islands, usa. *Journal of Taphonomy*, 1(1):43–66, 2003.

Kowalewski M., Serrano G.E.A., Flessa K.W., and Goodfriend G.A. Dead delta's former productivity: two trillion shells at the mouth of the colorado river. *Geology*, 28(12): 1059–1062, 2000. doi:10.1130/0091-7613200028<1059:DDFPTT<2.0.CO;2.

Krug A.Z., Jablonski D., and Valentine J.W. Species–genus ratios reflect a global history of diversification and range expansion in marine bivalves. *Proceedings of the Royal Society B*, 275:1117–1123, 2008. doi:10.1098/rspb.2007.1729.

Kruskal J.B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29(1):1–27, 1964. doi:10.1007/BF02289565.

Kusnerik K.M., Means G.H., Portell R.W., Brenner M., Hua Q., Kannai A., Means R., Monroe M.A., and Kowalewski M. Live, dead, and fossil mollusks in florida freshwater springs and spring-fed rivers: taphonomic pathways and the formation of multisourced, time-averaged death assemblages. *Paleobiology*, 46(3):356–378, 2020. doi:10.1017/pab.2020.25.

LACSD (Los Angeles County Sanitation Districts). *Joint Water Pollution Control Plant Biennial Receiving Water Monitoring Report 2010-2011*. Los Angeles County Sanitation Districts, Whittier, CA, 2020.

LACSD. *Joint Water Pollution Control Plant Biennial Receiving Water Monitoring Report 2018-2019*. Los Angeles County Sanitation Districts, Whittier, CA, 2020.

LACSD. *Joint Water Pollution Control Plant Biennial Receiving Water Monitoring Report 2020-2021*. Los Angeles County Sanitation Districts, Whittier, CA, 2022.

Landeiro V.L., Bini L.M., Costa F.R.C., Franklin E., Nogueira A., de Souza J.L.P., Moraes J., and Magnusson W.E. How far can we go in simplifying biomonitoring assessments? an integrated analysis of taxonomic surrogacy, taxonomic sufficiency and numerical resolution in a megadiverse region. *Ecological Indicators*, 23:366–373, 2012. doi:10.1016/j.ecolind.2012.04.023.

LeClaire A.M., Powell E.N., Mann R., Hemeon K.M., Pace S.M., Sower J.R., and Redmond T.E. Historical biogeographic range shifts and the influence of climate change on ocean quahogs (arctica islandica) on the mid-atlantic bight. *The Holocene*, 32(9):964–976, 2022. doi:10.1177/09596836221101275.

Lee H.J., Sherwood C.R., Drake D.E., Edwards B.D., Wong F., and Hamer M. Spatial and temporal distribution of contaminated, effluent-affected sediment on the palos verdes margin, southern california. *Continental Shelf Research*, 22(6-7):859–880, 2002. doi:10.1016/S0278-4343(01)00108-X.

Leonard-Pingel J.S., Kidwell S.M., Tomašových A., Alexander C.R., and Cadien D.B. Gauging benthic recovery from 20th century pollution on the southern california continental shelf using bivalves from sediment cores. *Marine Ecology Progress Series*, 615:101–119, 2019. doi:10.3354/meps12918.

Leshno Y., Benjamini C., and Edelman-Furstenberg Y. Ecological quality assessment in the eastern mediterranean combining live and dead molluscan assemblages. *Marine Pollution Bulletin*, 104(1-2):246–256, 2016. doi:10.1016/j.marpolbul.2016.01.014.

Leshno Y., Edelman-Furstenberg Y., Mienis H., and Benjamini C. Molluscan live and dead assemblages in an anthropogenically stressed shallow-shelf: Levantine margin of israel. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 433:49–59, 2015. doi:10.1016/j.palaeo.2015.05.008.

Lie U. and Evans R.A. Long-term variability in the structure of subtidal benthic communities in puget sound, washington, usa. *Marine Biology*, 21:122–126, 1973. doi:10.1007/BF00354608.

Lim J.Y., Tay T.S., Lim C.S., Lee S.S.C., Teo S.L.-M., and Tan K.S. Mytella strigata (bivalvia: Mytilidae): an alien mussel recently introduced to singapore and spreading rapidly. *Molluscan Research*, 38(3):170–186, 2018. doi:10.1080/13235818.2018.1423858.

Lloyd G.T., Young J.R., and Smith A.B. Taxonomic structure of the fossil record is shaped by sampling bias. *Systematic Biology*, 61(1):80, 2012. doi:10.1093/sysbio/syr076.

Lodeiros C., Hernández-Reyes D., Salazar J.M., Rey-Méndez M., and González-Henríquez N. First report of the mussel mytella strigata (hanley, 1843) in the venezuelan caribbean from an invasion in a shrimp farm. *Latin American Journal of Aquatic Research*, 49(3): 531–537, 2021a. doi:10.3856/vol49-issue3-fulltext-2626.

Long E.R. and Chapman P.M. A sediment quality triad: measures of sediment contamination, toxicity and infaunal community composition in puget sound. *Marine Pollution Bulletin*, 16(10):405–415, 1985. doi:10.1016/0025-326X(85)90290-5.

Lu X., Xu J., Xu Z., and Liu X. Assessment of benthic ecological quality status using mulit-biotic indices based on macrofaunal assemblages in a semi-enclosed bay. *Frontiers in Marine Science*, 8:734710, 2021. doi:10.3389/fmars.2021.734710.

MacCready P., McCabe R.M., Siedlecki S.A., Lorenz M., Giddings S.N., Bos J., Albertson S., Banas N.S., and Garnier S. Estuarine circulation, mixing, and residence times in the salish sea. *Journal of Geophysical Research: Oceans*, 126(2):e2020JC016738, 2021. doi:10.1029/2020JC016738.

Macdonald T.A., Burd B.J., Macdonald V.I., and van Roodselaar A. *Taxonomic and feeding guild classification for the marine benthic macroinvertebrates of the Strait of Georgia, British Columbia.* Canadian Technical Report of Fisheries and Aquatic Sciences 2874, 2010.

Magierowski R.H. and Johnson C.R. Robustness of surrogates of biodiversity in marine benthic communities. *Ecological Applications*, 16(6):2264–2275, 2006. doi:10.1890/1051-0761(2006)016[2264:ROSOBI]2.0.CO;2.

Mantel N. The detection of disease clustering and a generalized regression approach. *Cancer Research*, 27(2Part1):209–220, 1967. doi:10.1029/2020JC016738.

Martinelli J.C., Madin J.S., and Kosnik M.A. Dead shell assemblages faithfully record living molluscan assemblages at one tree reef. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 457:158–169, 2016. doi:10.1016/j.palaeo.2016.06.002.

Martínez-Crego B., Alcoverro T., and Romero J. Biotic indices for assessing the status of coastal waters: a review of strengths and weaknesses. *Journal of Environmental Monitoring*, 12(5):1013–1028, 2010. doi:10.1039/b920937a.

Maurer D., Nguyen H., Robertson G., and Gerlinger T. The infaunal trophic index (iti): its suitability for marine environmental monitoring. *Ecological Applications*, 9(2):699–713, 1999. doi:10.1890/1051-0761(1999)009[0699:TITIII]2.0.CO;2.

Meadows C.A., Grebmeier J.M., and Kidwell S.M. High-latitude benthic bivalve biomass and recent climate change: testing the power of live-dead discordance in the pacific arctic. *Deep Sea Research Part II: Topical Studies in Oceanography*, 162:152–163, 2019. doi:10.1016/j.dsr2.2019.04.005.

Meadows C.A., Grebmeier J.M., and Kidwell S.M. Arctic bivalve dead-shell assemblages as high temporal- and spatial- resolution archives of ecological regime change in response to climate change. *Geological Society, London, Special Publications*, 529(1):99–130, 2023. doi:10.1144/SP529-2022-131.

Meldahl K.H., Flessa K.W., and Cutler A.H. Time-averaging and postmortem skeletal survival in benthic fossil assemblages: quantitative comparisons among holocene environments. *Paleobiology*, 23(2):207–229, 1997. doi:10.1017/S0094837300016791.

Mellin C., Delean S., Caley J., Edgar G., and Meekan M. Effectiveness of biological surrogates for predicting patterns of marine biodiversity: a global meta-analysis. *PLOS ONE*, 6(6): e20141, 2011. doi:10.1371/journal.pone.0020141.

Mistri M. and Munari C. The performance of biomass-based ambi in lagoonal ecosystems. *Marine Pollution Bulletin*, 99(1-2):126–137, 2015. doi:10.1016/j.marpolbul.2015.07.048.

Mitchem E.L., Fajans J.S., and Baker S.M. Contrasting responses of two native crustaceans to nonindigenous prey, the green mussel, perna viridis. *Florida Scientist*, 70(2):180–188, 2007.

Moore S.K., Mantua N.J., Newton J.A., Kawase M., Warner M.J., and Kellogg J.P. A descriptive analysis of temporal and spatial patterns of variability in puget sound oceanographic properties. *Estuarine, Coastal and Shelf Science*, 80(4):545–554, 2008. doi:10.1016/j.ecss.2008.09.016.

Moreno C.E., Sanchez-Rojas G., Pineda E., and Escobar F. Shortcuts for biodiversity evaluation: a review of terminology and recommendations for the use of target groups, bioindicators and surrogates. *International Journal of Environmental Health*, 1(1):71–86, 2007. doi:10.1504/IJENVH.2007.012225.

Moritz C., Brandl S.J., Rouzé H., Vii J., Pérez-Rosales G., Bosserelle P., Chancerelle Y., Galzin R., Liao V., Siu G., Taiarui M., Nugues M.M., and Hédouin L. Long-term monitoring of benthic communities reveals spatial determinants of disturbance and recovery dynamics on coral reefs. *Marine Ecology Progress Series*, 672:141–152, 2021b. doi:10.3354/meps13807.

Mueller M., Pander J., and Geist J. Taxonomic sufficiency in freshwater ecosystems: effects of taxonomic resolution, functional traits, and data transformation. *Freshwater Science*, 32(3):762–778, 2013. doi:10.1899/12-212.1.

Muxika I., Borja Á., and Bald J. Using historical data, expert judgement and multivariate analysis in assessing reference conditions and benthic ecological status, according to the european water framework directive. *Marine Pollution Bulletin*, 55(1-6):16–29, 2007. doi:10.1016/j.marpolbul.2006.05.025.

Muxika I., P.J. Somerfield, Borja Á., and Warwick R.M. Assessing proposed modifications to the atzi marine biotic index (ambi), using biomass and production. *Ecological Indicators*, 12(1):96–104, 2012. doi:10.1016/j.ecolind.2011.04.030.

Nawrot R., Berensmeier M., Gallmetzer I., Haselmair A., Tomašových A., and Zuschin M. Multiple phyla, one time resolution? similar time averaging in benthic foraminifera, mollusk, echinoid, crustacean, and otolith fossil assemblages. *Geology*, 50(8):902–906, 2022. doi:10.1130/G49970.1.

Nerlović V., Doğan A., and Hrs-Brenko M. Response to oxygen deficiency (depletion): bivalve assemblages as an indicator of ecosystem instability in the northern adriatic sea. *Biologia*, 66(6):1114–1126, 2011. doi:10.2478/s11756-011-0121-3.

Nichols F.H. Interdecadal change in the deep puget sound benthos. *Hydrobiologia*, 493: 95–114, 2003. doi:10.1023/A:1025453700512.

Norkko J., Pilditch C.A., gammal J., Rosenberg R., Enemar A., Magnusson M., Granberg M.E., Lindgren J.F., Agrenius S., and Norkko A. Ecosystem functioning along gradients of increasing hypoxia and changing soft-sediment community types. *Journal of Sea Research*, 153:101781, 2019. doi:10.1016/j.seares.2019.101781.

OCSD (Orange County Sanitation Districts. *Marine Monitoring Annual Report 2021-2022*. Orange County Sanitation Districts, Fountain Valley, CA, 2023.

Olsgard F., Brattegard T., and Holthe T. Polychaetes as surrogates for marine biodiversity: lower taxonomic resolution and indicator groups. *Biodiversity and Conservation*, 12:1033–1049, 2003. doi:10.1023/A:1022800405253.

Olsgard F. and Somerfield P.J. Surrogates in marine benthic investigations - which taxonomic unit to target? *Journal of Aquatic Ecosystem Stress and Recovery*, 7:25–42, 2000. doi:10.1023/A:1009967313147.

Olszewski T.D. A unified mathematical framework for the measurement of richness and evenness within and among multiple communities. *Oikos*, 104(2):377–387, 2004. doi:10.1111/j.0030-1299.2004.12519.x.

Olszewski T.D. and Kaufman D.S. Tracing burial history and sediment recycling in a shallow estuarine setting(copano bay, texas) using postmortem ages of the bivalve mulinia lateralis. *Palaios*, 30(3):224–237, 2015. doi:10.2110/palo.2014.063.

Omena E.P., Lavrado H.P., Paranhos R., and Silva T.A. Spatial distribution of intertidal sandy beach polychaeta along an estuarine and morphodynamic gradient

in an eutrophic tropical bay. *Marine Pollution Bulletin*, 64(9):1861–1873, 2012. doi:10.1016/j.marpolbul.2012.06.009.

Pandolfi J.M. Numerical and taxonomic scale of analysis in paleoecological data sets: examples from neo-tropical pleistocene reef coral communities. *Journal of Paleontology*, 75 (3):546–563, 2001. doi:10.1666/0022-3360(2001)075<0546:NATSOA>2.0.CO;2.

Partridge V., Weakland S., Dutch M., Burgess D., and Eagleston A. *Sediment quality in Puget Sound: changes in chemical contaminants and invertebrate communities at 10 sentinel stations, 1989-2015*. Washington State Department of Ecology Publication, Lacey, WA, 2018.

Pauly D. Anecdotes and the shifting baseline syndrome of fisheries. *Trends in Ecology and Evolution*, 10(10):430, 1995. doi:10.1016/s0169-5347(00)89171-5.

Pearson T.H. and Rosenberg R. Macrobenthic succession in relation to organic enrichment and pollution of the marine environment. *Oceanography and Marine Biology: an Annual Review*, 16:229–311, 1978.

Pelletier M.C. and Charpentier M. Assessing the relative importance of stressors to the benthic index, m-ambi: an example from u.s. estuaries. *Marine Pollution Bulletin*, 186: 114456, 2023. doi:10.1016/j.marpolbul.2022.114456.

Pelletier M.C., Gillett D.J., Hamilton A., Grayson T., Hansen V., Leppo E.W., Weisberg S.B., and Borja A. Adaptation and application of multivariate ambi (m-ambi) in us coastal waters. *Ecological Indicators*, 89:818–827, 2018. doi:10.1016/j.ecolind.2017.08.067.

Peterson C.H. Relative abundances of living and dead molluscs in two californian lagoons. *Lethaia*, 9:137–148, 1976. doi:10.1111/j.1502-3931.1976.tb00958.x.

Pik A.J., Oliver I., and Beattie A.J. Taxonomic sufficiency in ecological studies of terrestrial invertebrates. *Australian Journal of Ecology*, 24(5):555–562, 1999. doi:10.1046/j.1442-9993.1999.01003.x.

Pitacco V., Mistri M., Aleffi I.F., Lardicci C., Prato S., Tagliapietra D., and Munari C. The efficiency of taxonomic sufficiency for identification of spatial patterns at different scales in transitional waters. *Marine Environmental Research*, 144:84–91, 2019. doi:10.1016/j.marenvres.2019.01.001.

Powell E.N., Kuykendall K.M., and Moreno P. The death assemblage as a marker for habitat and an indicator of climate change: Georges bank surfclams and open quahogs. *Continental Shelf Research*, 142:14–31, 2017. doi:10.1016/j.csr.2017.05.008.

Powell E. and Stanton R.J. Jr. Estimating biomass and energy flow of mollusks in palaeocommunities. *Paleontology*, 28:1–34, 1985.

Power A.J., Walker R.L., Payne K., and Hurley D. First occurrence of the nonindigenous green mussel, perna viridis (linnaeus, 1758) in coastal georgia, united states. *Journal of Shellfish Research*, 23(3):741–745, 2004.

Prior K., Adams D.C., Klepzig K., and Hulcr J. When does invasive species removal lead to ecological recovery? implications for management success. *Biological Invasions*, 20: 267–283, 2018. doi:10.1007/s10530-017-1542-x.

Pruden M.J., Dietl G.P., Handley J.C., and Smith J.A. Using molluscs to assess ecological quality status of soft-bottom habitats along the atlantic coastline of the united states. *Ecological Indicators*, 129:107910, 2021. doi:10.1016/j.ecolind.2021.107910.

PSEP (Puget Sound Estuary Program). *Recommended protocols for sampling and analyzing subtidal benthic macroinvertebrate assemblages in Puget Sound*. Tetra Tech Inc., Bellevue, WA, 1987.

R Core Team. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, 2023.

Ranasinghe J.A., Stein E.D., Frazier M.R., and Gillett D.J. *Development of Puget Sound benthic indicators*. Southern California Coastal Water Research Project Technical Report 755, Costa Mesa, CA, 2013.

Rice M.A., Rawson P.D., Salinas A.D., and Rosario W.R. Identification and salinity tolerance of the western hemisphere mussel mytella charruana (d'orbigny, 1842) in the philippines. *Journal of Shellfish Research*, 35(4):865–873, 2016. doi:10.2983/035.035.0415.

Riedel B., Zuschin M., and Stachowitsch M. Tolerance of benthic macrofauna to hypoxia and anoxia in shallow coastal seas: a realistic scenario. *Marine Ecology Progress Series*, 458:39–52, 2012. doi:10.3354/meps09724.

Rittenberg S.C., Mittwer T., and Iyler D. Coliform bacteria in sediments around three marine sewage outfalls. *Limnology and Oceanography*, 3(1):101–108, 1958. doi:10.4319/lo.1958.3.1.0101.

Ritter M.D.N., Erthal F., Kosnik M.A., Coimbra J.C., and Kaufman D.S. Spatial variation in the temporal resolution of subtropical shallow-water mollsucan death assemblages. *Palaios*, 32(9):572–583, 2017. doi:10.2110/palo.2017.003.

Rojas A., Calatayud J., Kowalewski M., Neuman M., and Rosvall M. A multiscale view of the phanerozoic fossil record reveals the three major biotic transitions. *Communications Biology*, 4:309, 2021b. doi:10.1038/s42003-021-01805-y.

Rose D. and Webber M. Characterization of microplastics in the surface waters of kingston harbour. *Science of the Total Environment*, 664:753–760, 2019. doi:10.1016/j.scitotenv.2019.01.319.

Rosenberg R., Blomqvist M., Nilsson H.C., Cederwall H., and Dimming A. Marine quality assessment by use of benthic species-abundance distributions: a proposed new protocol within the european union water framework directive. *Marine Pollution Bulletin*, 49(9-10): 728–739, 2004a. doi:10.1016/j.marpolbul.2004.05.013.

Ruppert K.M., Kline R.J., and Rahman M.S. Past, present, and future perspectives of environmental dna (edna) metabarcoding: a systematic review in methods, monitoring, and applications of global edna. *Global Ecology and Conservation*, 17:e00547, 2019. doi:10.1016/j.gecco.2019.e00547.

Rylander K., Perez J., and Gomez J. The distribution of the brown mussel, perna perna and the green mussel p. viridis (mollusca:bivalvia:mytilidae) in north eastern venezuela. *Caribbean Marine Studies*, 5:86–87, 1996.

Sander L., Hass H.C., Michaelis R., Groß C., Hausen T., and Pogoda B. The late holocene demise of a sublittoral oyster bed in the north sea. *PLOS ONE*, 16(2):e0242208, 2021. doi:10.1371/journal.pone.0242208.

Sanpanich K. and Wells F.E. Mytella strigata (hanley, 1843) emerging as an invasive marine threat in southeast asia. *BioInvasions Records*, 8(2):343–356, 2019. doi:10.3391/bir.2019.8.2.16.

SCAMIT (Southern California Association of Marine Invertebrate Taxonomists). *A Taxonomic Listing of Benthic Macro- and Megainvertebrates from Infaunal and Epifaunal Monitoring and Research Programs in the Southern California Bight, 13th Edn.* Southern California Association of Marine Invertebrate Taxonomists, Los Angeles, CA, 2021.

Schiff K., Greenstein D., Dodder N., and Gillett D.J. Sourthern california bight regional monitoring. *Regional Studies in Marine Science*, 4:34–46, 2016. doi:10.1016/j.rsma.2015.09.003.

Schopf T.J.M. Fossilization potential of an intertidal fauna: Friday harbor, washington. *Paleobiology*, 4(3):261–270, 1978. doi:10.1017/S0094837300005996.

Schwing P.T., O'Malley B.J., Romero I.C., Martinez-Colón M., Hastings D.W., Glabach M.A., Hladky E.M., Greco A., and Hollander D.J. Characterizing the variability of benthic foraminifera in the northeastern gulf of mexico following the deepwater horizon event (2010–2012). *Environmental Science and Pollution Research*, 24:2754–2769, 2017. doi:10.1007/s11356-016-7996-z.

Schopf T.J.M. A factor analytic description of the phanerozoic marine fossil record. *Paleobiology*, 7(1):36–53, 1981. doi:10.1017/S0094837300003778.

Sherwin T.J. and Deeming K.R. *Water Circulation and its Relation to Pollution in Kingston Harbour, Jamaica.* Project UCES 8 Report U80-1, 1980.

Shiganova T.A., Sommer U., Javidpour J., Molinero J.C., Malej A., Kazmin A.S., Isinibilir M., Christou E., Siokou-Frangou I., Marambio M., V. Fuentes, Mirsoyan Z.A., Gülsahin N., Lombard F., Lilley M.K.S., Angel D.L., Galil B.S., Bonnet D., and Delpy F. Patterns of invasive ctenophore mnemiopsis leidyi distribution and variability in different recipient environments of the eurasian seas: a review. *Marine Environmental Research*, 152:104791, 2019. doi:10.1016/j.marenvres.2019.104791.

Siddall S.E. A clarification of the genus perna (mytilidae). *Bulletin of Marine Science*, 30 (4):858–870, 1980.

Simonini R., Ansaloni I., Bonvicini Pagliai A.M., and Prevedelli D. Organic enrichment and structure of the macrozoobenthic community in the northern adriatic sea in an area facing adige and po mouths. *ICES Journal of Marine Science*, 61(6):871–881, 2004b. doi:10.1016/j.icesjms.2004.06.018.

Siung A.M. Studies on the biology of isognomon alatus gmelin (bivalvia: Isognomonidae) with notes on its potential as a commercial species. *Bulletin of Marine Science*, 30(1): 90–101, 1980.

Smith J.A. and Dietl G.P. The value of geohistorical data in identifying a recent human-induced range expansion of a predatory gastropod in the colorado river delta, mexico. *Journal of Biogeography*, 43(4):791–800, 2016. doi:10.1111/jbi.12644.

Smith J.A., Pruden M.J., Handley J.C., Durham S.R., and Dietl G.P. Assessing the utility of death assemblages as reference conditions in a common benthic index (m-ambi) with simulations. *Geological Society, London, Special Publications*, 529(1):131–151, 2023. doi:10.1144/SP529-2022-89.

Smith R.W., Bergen M., Weisberg S.B., Cadien D., Dalkey A., Montagne D., Stull J.K., and Velarde R.G. Benthic response index for assessing infaunal communities on the southern california mainland shelf. *Ecological Applications*, 11(4):1073–1087, 2001. doi:10.1890/1051-0761(2001)011[1073:BRIFAI]2.0.CO;2.

Smith R.W., Ranasinghe J.A., Weisberg S.B., Montagne D.E., Cadien D.B., Mikel T.K., Velarde R.G., and Dalkey A. *Extending the Southern California Benthic Response Index to Assess Benthic Condition in Bays*. Southern California Coastal Water Research Project Technical Report 410, Westminster, CA, 2003.

Staff G.M. and Powell E.N. Taxonomic levels, in marine community studies, revisited. *Marine Ecology Progress Series*, 127:113–119, 1995. doi:10.3354/meps127113.

Souza J.L.P., Baccaro F.B., Landeiro V.L., Franklin E., Magnusson W.E., Pequeno P.A.C.L., and Fernandes I.O. Taxonomic sufficiency and indicator taxa reduce sampling costs and increase monitoring effectiveness for ants. *Diversity and Distributions*, 22(1):111–122, 2016. doi:10.1111/ddi.12371.

St-Pierre A.P., Shikon V., and Schneider D.C. Count data in biology – data transformation or model reformation? *Ecology and Evolution*, 8(6):3077–3085, 2018. doi:10.1002/ece3.3807.

Staff G.M. and Powell E.N. The paleoecological significance of diversity: the effect of time averaging and differential preservation on macroinvertebrate species richness in death assemblages. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 63(1-3):73–89, 1988. doi:10.1016/0031-0182(88)90091-0.

Staff G., Powell E.N., Stanton R.J. Jr., and Cummins H. Biomass: is it a useful tool in paleocommunity reconstruction? *Lethaia*, 18(3):209–232, 1985. doi:10.1111/j.1502-3931.1985.tb00700.

Staff G.M., Stanton R.J. Jr., Powell E., and Cummins H. Time-averaging, taphonomy, and their impact on paleocommunity reconstruction: death assemblages in texas bays. *Geological Society of America Bulletin*, 97:428–443, 1986.

Stanley S.M. *Relation of shell form to life habits of the Bivalvia (Mollusca)*. Geological Society of America Memoir 125, Boulder, CO, 1970.

Stark J.S., Kim S.L., and Oliver J.S. Anthropogenic disturbance and biodiversity of marine benthic communities in antarctica: a regional comparison. *PLOS ONE*, 9:e98802, 2014. doi:10.1371/journal.pone.0098802.

Steger J., Bošnjak M., Belmaker J., Galil B.S., Zuschin M., and Albano P.G. Non-indigenous molluscs in the eastern mediterranean have distinct traits and cannot replace historic ecosystem functioning. *Global Ecology and Biogeography*, 31(1):89–102, 2022. doi:10.1111/geb.13415.

Stein E.D. and Cadien D.B. Ecosystem response to regulatory and management actions: the southern california experience in long-term monitoring. *Marine Pollution Bulletin*, 59: 91–100, 2009. doi:10.1016/j.marpolbul.2009.02.025.

Strayer D.L. Twenty years of zebra mussels: lessons from the mollusk that made headlines. *Frontiers in Ecology and the Environment*, 7(3):135–141, 2009. doi:10.1111/jbi.12644.

Strayer D.L., D'Antonio C.M., Essl F., Fowler M.S., Geist J., Hilt S., Jarić I., Jöhnk K., Jones C.G., Lambin X., Latzka A.W., Pergl J., Pyšek P., Robertson P., von Schmalensee M., Stefansson R.A., Wright J., and Jeschke J.M. Boom-bust dynamics in biological invasions: towards an improved application of the concept. *Ecology Letters*, 20(10):1337–1350, 2017. doi:10.1111/ele.12822.

Stull J.K., D.J. Swift, and Niedoroda A.W. Contaminant dispersal on the palos verdes continental margin. i. sediments and biota near a major california wastewater discharge. *Science of the Total Environment*, 179:73–90, 1996. doi:10.1016/S0048-9697(96)90050-9.

Swartz R.C., Cole F.A., Schults D.W., and DeBen W.A. Ecological changes in the southern california bight near a large sewage outfall: benthic conditions in 1980 and 1983. *Marine Ecology Progress Series*, 31(1):1–13, 1986.

Swartz R.C., Schults D.W., Lamberson J.O., Ozretich R.J., and Stull J.K. Vertical profiles of toxicity, organic carbon, and chemical contaminants in sediment cores from the palos verdes shelf and santa monica bay, california. *Marine Environmental Research*, 31(3): 215–225, 1991. doi:10.1016/0141-1136(91)90012-W.

Tabak M.A., Piaggio A.J., Miller R.S., Sweitzer R.A., and Ernest H.B. Anthropogenic factors predict movement of an invasive species. *Ecosphere*, 8(6):e01844, 2017. doi:10.1002/ecs2.1844.

Thorne R.S.J., W.P. Williams, and Cao Y. The influence of data transformations on biological monitoring studies using macroinvertebrates. *Water Research*, 33(2):343–350, 1999. doi:10.1016/S0043-1354(98)00247-4.

Timms L.L., Bowden J.J., Summerville K.S., and Buddle C.M. Does species-level resolution matter? taxonomic sufficiency in terrestrial arthropod biodiversity studies. *Insect Conservation and Diversity*, 6(4):453–462, 2013. doi:10.1111/icad.12004.

Todd P.A., Heery E.C., Loke L.H.L., Thurstan R.H., Kotze D.J., and Swan C. Towards an urban marine ecology: characterizing the drivers, patterns and processes of marine ecosystems in coastal cities. *Oikos*, 128(9):1215–1242, 2019. doi:10.1111/oik.05946.

Tomašových A., Albano P.G., Fuksi T., Gallmetzer I., Haselmair A., Kowalewski M., Nawrot R., Nerlović V., Scarponi D., and Zuschin M. Ecological regime shift preserved in the anthropocene stratigraphic record. *Proceedings of the Royal Society B*, 287:20200695, 2020. doi:10.1098/rspb.2020.0695.

Tomašových A. and Kidwell S.M. Preservation of spatial and environmental gradients by death assemblages. *Paleobiology*, 35(1):119–145, 2009. doi:10.1666/07081.1.

Tomašových A. and Kidwell S.M. The effects of temporal resolution on species turnover and on testing metacommunity models. *The American Naturalist*, 175(5):587–606, 2010. doi:10.1086/651661.

Tomašových A. and Kidwell S.M. Nineteenth-century collapse of a benthic marine ecosystem on the open continental shelf. *Proceedings of the Royal Society B*, 284:20170328, 2017. doi:10.1098/rspb.2017.0328.

Tomašových A., Kidwell S.M., Alexander C.R., and Kaufman D.S. Millennial-scale age offsets within fossil assemblages: result of bioturbation below the taphonomic active zone and out-of-phase production. *Paleoceanography and Paleoclimatology*, 34(6):954–977, 2019b. doi:10.1029/2018PA003553.

Tomašových A., Kidwell S.M., and Barber R.F. Inferring skeletal production from time-averaged assemblages: skeletal loss pulls the timing of production pulses towards the modern period. *Paleobiology*, 42(1):54–76, 2016. doi:10.1017/pab.2015.30.

Tomašových A., Kidwell S.M., Barber R.F., and Kaufman D.S. Long-term accumulation of carbonate shells reflects a 100-fold drop in loss rate. *Geology*, 42(9):819–822, 2014b. doi:10.1130/G35694.1.

Tomašových A., Kidwell S.M., and Dai R. A downcore increase in time averaging is the null expectation from the transit of death assemblages through a mixed layer. *Paleobiology*, 49 (3):1–36, 2023. doi:10.1017/pab.2022.42.

Tweedley J.R., Warwick R.M., Clarke K.R., and Potter I.C. Family-level ambi is valid for use in the north-eastern atlantic but not for assessing the health of microtidal australian estuaries. *Estuarine, Coastal and Shelf Science*, 141:85–96, 2014. doi:10.1016/j.ecss.2014.03.002.

Tweitmann A. and Dietl G.P. Live-dead mismatch of molluscan assemblages indicates disturbance from anthropogenic eutrophication in the barnegat bay-little egg harbor estuary. *Journal of Shellfish Research*, 37(3):615–624, 2018. doi:10.2983/035.037.0314.

Tyler C.L. and Kowalewski M. Surrogate taxa and fossils as reliable proxies of spatial biodiversity patterns in marine benthic communities. *Proceedings of the Royal Society B*, 284:20162839, 2017. doi:10.1098/rspb.2016.2839.

Underwood A.J. Beyond baci: experimental designs for detecting human environmental impacts on temporal variations in natural populations. *Marine and Freshwater Research*, 42(5):569–587, 1991. doi:10.1071/MF9910569.

Underwood A.J. On beyond baci: sampling designs that might reliably detect environmental disturbances. *Ecological Applications*, 4(1):3–15, 1994. doi:10.2307/1942110.

Van Hoey G., Borja A., Birchenough S., Buhl-Mortensen L., Degraer S., Fleischer D., Kerckhof F., Magni P., Muxika I., Reiss H., Schröder A., and M.L. Zettler. The use of benthic indicators in europe: from the water framework directive to the marine strategy framework directive. *Marine Pollution Bulletin*, 60(12):2187–2196, 2010. doi:10.1016/j.marpolbul.2010.09.015.

Vijapure T. and Sukumaran S. Optimization of the taxonomic resolution of an indicator taxon for cost-effective ecological monitoring: perspectives from a heterogeneous tropical coastline. *Journal of Environmental Management*, 247:474–483, 2019. doi:10.1016/j.jenvman.2019.05.154.

Villnäs A. and Norkko A. Benthic diversity gradients and shifting baselines: implications for assessing environmental status. *Ecological Applications*, 21(6):2172–2186, 2011. doi:10.1890/10-1473.1.

Villnäs A., Norkko A., Lukkari K., Hewitt J., and Norkko A. Consequences of increasing hypoxic disturbance on benthic communities and ecosystem functioning. *PLOS ONE*, 7: e44920, 2012. doi:10.1371/journal.pone.0044920.

Wade B.A., Antonio L., and Mahon R. Increasing organic pollution in kingston harbour, jamaica. *Marine Pollution Bulletin*, 7(3):106–111, 1972. doi:10.1016/0025-326X(72)90251-2.

Waldbusser G.G., Powell E.N., and Mann R. Ecosystem effects of shell aggregations and cycling in coastal waters: an example of chesapeake bay oyster reefs. *Ecology*, 94(4): 895–903, 2013. doi:10.1890/12-1179.1.

Warme J.E. Live and dead molluscs in a coastal lagoon. *Journal of Paleontology*, 43(1): 141–150, 1969.

Warwick R.M. The level of taxonomic discrimination required to detect pollution effects on marine benthic communities. *Marine Pollution Bulletin*, 19(6):259–268, 1988. doi:10.1016/0025-326X(88)90596-6.

Warwick R.M. Environmental impact studies on marine communities: pragmatical considerations. *Australian Journal of Ecology*, 18(1):63–80, 1993. doi:10.1111/j.1442-9993.1993.tb00435.x.

Warwick R.M. and Clarke K.R. Relearning the abc: taxonomic changes and abundance/biomass relationships in disturbed benthic communities. *Marine Biology*, 118:739–744, 1994. doi:10.1007/BF00347523.

Warwick R.M., Pearson T.H., and Ruswahyuni. Detection of pollution effects on marine macrobenthos: further evaluation of the species abundance/biomass method. *Marine Biology*, 95:193–200, 1987. doi:10.1007/BF00409005.

Weakland S., Partridge V., and Dutch M. *Sediment quality in Puget Sound: changes in chemistry, toxicity, and benthic invertebrates at multiple geographic scales, 1989-2015*. Washington State Department of Ecology Publication, Lacey, WA, 2018.

Webber M.K., Webber D.F., Ranston E.R., Dunbar F.N., and Simmonds R.M.A. Detection of pollution effects on marine macrobenthos: further evaluation of the species abundance/biomass method. *Bulletin of Marine Science*, 73(2):361–378, 2003.

Westgate M.J., Barton P.S., Lane P.W., and D.B. Lindenmayer. Global meta-analysis reveals low consistency of biodiversity congruence relationships. *Nature Communications*, 5(1): 3899, 2014. doi:10.1038/ncomms4899.

Włodarska-Kowalczuk M. and Kędra M. Surrogacy in natural patterns of benthic distribution and diversity: selected taxa versus lower taxonomic resolution. *Marine Ecology Progress Series*, 351:53–63, 2007. doi:10.3354/meps07127.

Yanes Y. Anthropogenic effect recorded in the live-dead compositional fidelity of land snail assemblages from san salvador island, bahamas. *Biodiversity and Conservation*, 21:3445–3466, 2012. doi:10.1007/s10531-012-0373-4.

Zuschin M., Nawrot R., Harzhauser M., Mandic O., and Tomašových A. Taxonomic and numerical sufficiency in depth- and salinity-controlled marine paleocommunities. *Paleobiology*, 43(3):463–478, 2017. doi:10.1017/pab.2016.49.