

The use of diversity indices for local assessment of marine sediment quality

Shinya Hosokawa (✉ hosokawa@p.mpat.go.jp)

Port and Airport Research Institute

Kyosuke Momota

Marine Ecology Research Institute

Anthony Chariton

Macquarie University

Ryoji Naito

National Institute for Land and Infrastructure Management

Yoshiyuki Nakamura

Yokohama National University

Research Article

Keywords: community structure, species density, Shannon index, biological communities

DOI: <https://doi.org/10.21203/rs.3.rs-279143/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Diversity indices are commonly used to measure changes in marine benthic communities. Although diversity indices give a good understanding of how biological communities respond to changes in sediment, responses of these indices can vary because of small sample sizes, and insensitive responses can result from an inappropriate choice of test community. This study focused on species density and two indices of community structure and explored how they respond to variations in sediment characteristics in our target region, Japan, and in two local areas within this region. Our analysis of the Japanese regional dataset showed a decrease in species density and the dominance of a few species as sediment conditions degraded. Local case studies showed that species density responded to sediment degradation at sites where the community was variable. However, the indices for community structure sometimes became misleading because of inaccurate estimates with small sample sizes, and could become insensitive because of masking by community variability. We conclude that species density is a sensitive indicator of change for marine benthic communities, but that indices of community structure should only be used when there is a sufficient sample size and a test community is distinguishable from other coexisting communities.

Background

Diversity indices are used to measure the richness and evenness of species diversity. Because these indices aid in the interpretation of changes in benthic communities, they are used as ecological approaches for evaluating marine sediments¹ and have advantages over several other approaches, such as chemical and toxicological evaluations, because of their more realistic application under field conditions².

For marine benthic invertebrate communities, these indices are usually estimated from counts of individuals obtained by using areal unit samplers such as quadrats or bottom samplers. Species richness, the Shannon index, and Pielou evenness are diversity indices estimated from these variables, and they tend to decrease with increasing environmental contamination³. In addition, the areal species richness (species density), which normalizes the species richness in a community to the area sampled⁴, is also used to assess changes in benthic communities.

However, small sample sizes (i.e. small numbers of individuals sampled) can introduce considerable uncertainty. Although small sample size is not a problem when estimating species density because of the way the index is defined, it often causes uncertainty in other diversity indices. For example, species richness follows an increasing curve (called a rarefaction curve) as the number of reference individuals increases⁵. Indices for community structure, which become asymptotic at infinite sample size (i.e. for the true relative abundances⁶), have been shown to have rarefaction curves similar to those for species richness⁶. The exponential of the Shannon index and the inverse Simpson's concentration index are representative of other indices that show rarefaction curves. Pielou evenness is also understood to be an index for community structure, because it is calculated from the Shannon index. These indices can substantially underestimate the true values when the number of individuals observed is small⁷.

Assessments of sediment quality may serve to test whether diversity indices respond to changes in sediment quality in a community. However, marine invertebrate communities are often also affected by small-scale environmental gradients in other environmental factors^{8,9}. If an appropriate community for testing cannot be identified, real community responses can be masked by potential variabilities in coexisting communities^{1,3}. When the pattern of community variability is predictable, it may be relatively easy to isolate the target community and test its response to sediment contamination. However, because the spatial pattern of benthic communities is a complex mosaic in areas of marine urbanization¹⁰, it can be difficult to identify and distinguish between a test community and other communities.

From these perspectives of sample size and community variability, we wondered whether diversity indices could work as indicators for sediment quality assessment under real field conditions. To address this, we proposed two questions. First, although species richness, the Shannon index, and Pielou evenness are known to respond to sediment degradation³, we wondered whether their responses are real, considering their dependence on marine benthic invertebrates, which are usually in low abundance in degraded sediments. Second, we wondered whether these indices are sensitive to changes in marine benthic invertebrates at the local scale we want to assess usually.

In this study, we chose to focus on three diversity indices: taxonomic density, the inverse Simpson's concentration index, and Pielou evenness. Although the Simpson's concentration index is better known than its inverse, we chose the inverse index (also known as "the effective number of species") because of its meaning as community structure¹¹. We explored how to produce robust assessments of sediment quality by estimating these diversity indices with high confidence (i.e. with appropriate consideration of sample size and community variability) in our target region, Japan. Our investigation was performed at the regional scale and in two local areas at intermediate latitudes within this region (**Fig. 1**). The regional dataset has been analysed so that species density is related to a group of sediment variables, which include measurements of softness and contaminants¹². We reanalysed the sediment variables that could potentially impact taxonomic density in this region, and explored the impacts of these variables on the inverse Simpson's concentration index and Pielou evenness. In addition, the two local datasets were analysed to determine how the diversity indices responded to sediment variables at the local scale. Finally, we outline a strategy for the ecological assessment of sediment quality using diversity indices in local marine areas.

Results

Potential responses of diversity indices.

The Polychaeta were the most diverse class of benthic invertebrates in the regional dataset, and Malacostraca, Bivalvia, and Gastropoda were the next most diverse classes overall after Polychaeta (**Table 1**). The composition of benthic invertebrate assemblages was relatively more diverse in the low-frequency group (see Materials and Methods) than in the other groups in the regional dataset, whereas Polychaeta were dominant in the high-frequency group. The average number of families observed at each location in the regional dataset was 12.4 ± 8.2 (mean \pm standard deviation) with a range of 0–36 (**Table 2**). We set up low-, intermediate-, and high-frequency groups to have a similar number of families among groups.

The sediment characteristic water content (WC, %), ranging from 24.7% to 360%, had significant power to explain the variability in family density in the overall regional dataset (**Fig. 2a**; the result of averaged model based on Akaike information criterion (AIC), and see Supplementary **Table S1** for the accepted candidate models). Water depth, latitude, and sample size were not significant. Also, the effect of WC was significantly negative in the low-frequency group (Supplementary **Fig. S1**). In this group, the ratio of the effect size of WC to the standard deviation of random effects was more than double in absolute value the values in the other groups (**Table 2**).

The inverse Simpson's concentration index for the overall data from the regional dataset ($N = 65$) was 4.3 ± 3.1 (mean \pm standard deviation). Pielou evenness was 0.71 ± 0.21 , where Pielou evenness was defined ($N = 60$). In the unreliable data (see Materials and Methods for the definition of reliable and unreliable data), the inverse Simpson's concentration index was not saturated (Supplementary **Fig. S2**), and Pielou evenness tended to be greater than in the reliable data at the same family density (Supplementary **Fig. S2**). The inverse Simpson's concentration index and Pielou evenness became 4.8 ± 3.4 and 0.64 ± 0.18 ($N = 36$), respectively, in the reliable data, where WC ranged between 32.7% and 263%.

The effect size of WC became negative for the inverse Simpson's concentration index in the analysis of reliable data (**Fig. 2b**). Among models with low AIC values, the effect size of WC was more negative in models that used only the reliable data than in models that used both reliable and unreliable data (Supplementary **Table S2**). The confidence interval for Pielou evenness extended farther into the negative range in the analysis of reliable data (**Fig. 2c**). This extension was produced because low-AIC models including WC were selected for the reliable data, whereas the explanatory variables varied in analyses with small AICs for the overall data (Supplementary **Table S3**). Although the effect of WC did not differ significantly from zero, its shifts toward the negative mean that there was a greater possibility of selecting a model including WC when using the reliable data. Sample coverage increased significantly from 0.79 ± 0.28 in the overall data to 0.94 ± 0.04 in the set of reliable data ($N = 36$; $t = -4.1$, $P = 1.2 \times 10^{-4}$; Welch two sample t-test), and from 0.86 ± 0.16 to 0.94 ± 0.04 in observations where Pielou evenness was defined ($N = 36$; $t = -3.5$, $P = 7.8 \times 10^{-4}$).

Assessments at two local sites.

Community structure in Matsunaga Bay.

In Matsunaga Bay, Polychaeta, Malacostraca, Bivalvia, and Gastropoda were the most diverse classes of benthic invertebrates (**Table 1**), which was the same as in the regional dataset. The ranges of values for other sediment variables were similar to those in the regional dataset except for sediment temperature (Supplementary **Fig. S3**). Because this local dataset was acquired in winter (see **Appendix S1**), sediment temperature (13.9 ± 1.9 °C; mean \pm standard deviation) was lower than the temperature in the regional dataset (24.8 ± 3.5 °C; between 10.4 and 30.4 °C).

Cluster analysis based on Sørensen dissimilarity was unable to show any apparent separations for species compositions in Matsunaga Bay (**Fig. 3a**); species compositions were similar within the intertidal flat at the mouth of a small river (sample locations IF1–IF5) and within the one in the inner bay (IF6–IF10). By contrast, cluster analysis based on Euclidean distance after $\log(n + 1)$ transformation showed that the community structure at the mouth of the small river was quite different from that at other locations in the bay (**Fig. 3b**). Distance-based redundancy analysis (dbRDA) showed that the four explanatory variables selected explained 56.9% of total variance in the benthic invertebrate community. The variability in community structure correlated with median sediment particle size (D_{50} ; 31.6%, $P = 0.0001$) (**Fig. 3c**), which appears as a difference in structure between the community at the mouth of the small river and the other observations in the first dbRDA coordinate axis. The polychaete *Simplisetia erythraeensis* was the most dominant in the community at the mouth of the small river. The variation in community structure was correlated with temperature (13.0%, $P = 0.0004$), depth (8.0%, $P = 0.0044$), and TOC (4.4%, $P = 0.042$).

Taken together, the cluster analysis and dbRDA for Matsunaga Bay indicate that there is a specific community at the mouth of the small river (hereafter, the “river-mouth” community) with more individuals than at the other locations. The data were reliable for this river-mouth community but unreliable at the other 25 locations.

Community structure in Nagoya Port.

The sediment variables in Nagoya Port had ranges similar to those in the regional dataset (Supplementary **Fig. S3**). The range of sediment temperatures was within that in the regional dataset because both datasets were obtained in summer (Supplementary **Appendix S1**).

The cluster analysis based on Sørensen dissimilarity was performed for all data from Nagoya Port except for sites N5, N9, N10, and N12, where there were no individuals sampled. The resulting tree showed that at the higher levels, the Fujimae intertidal flat and sites N6 and N17 could be separated from observations at other locations (**Fig. 4a**). However, the data for N6 and N17 were classified as unreliable. This means that the Fujimae intertidal flat was a habitat type that differed in species composition from other reliable data. The cluster analysis based on log-transformed abundance showed that N8 and N20 were separate specific habitat types in the community structure (**Fig. 4b**). The second outcome of this analysis is that the habitat type of the Fujimae intertidal flat can be separated from the other 11 locations with reliable data. The dbRDA coordination resulted in four explanatory variables explaining 38.4% of total variance in the benthic invertebrate community in Nagoya Port. Salinity separated the community structures of the Fujimae intertidal flat, N8, and N20 from other locations (**Fig. 4c**), however, this explained only 11.1% of the variation ($P = 0.001$). The other three explanatory variables also had relatively low contributions to the variability in community structure: the molar carbon to nitrogen ratio (C/N; 11.2%, $P = 0.016$), WC (9.2%, $P = 0.0006$), and D_{50} (6.8%, $P = 0.039$).

Taken together, these analyses suggest that sites N8 and N20 had specific community structures in terms of abundance, and that the Fujimae intertidal flat represents a distinct cluster in terms of species composition and abundance. These sites can be judged to include minor communities (see Materials and Methods). Conversely, the remaining reliable observations (data from 11 sites) were relatively isolated.

Responses of diversity indices at two local sites.

Because the analysis of the regional dataset showed that WC was the sediment variable with the most impact, we analysed trends in the diversity indices at the two local sites with WC as an explanatory variable.

Species density decreased significantly with increasing WC in all data types at the two local sites (**Table 3**). The slope was similar in the analyses for both overall and only test communities in Matsunaga Bay (**Fig. 5a**). In Nagoya Port, although the response to

WC was significant for both data types, the slope was more moderate in species density in the overall data than in the data excluding minor communities (**Fig. 5b**).

Sample coverage was 0.96 ± 0.02 (mean \pm standard deviation) in the river-mouth community and 0.61 ± 0.24 at the other 25 locations in Matsunaga Bay. The inverse Simpson's concentration index was saturated in the river-mouth community but not saturated at the other 25 locations (Supplementary **Fig. S4**). At the other 25 locations, Pielou evenness tended to be high, approaching 1.0 (Supplementary **Fig. S4**). In Nagoya Port, the inverse Simpson's concentration index was almost saturated for the reliable data (Supplementary **Fig. S5**), where sample coverage was 0.95 ± 0.04 . Pielou evenness was lower at the two sites with unreliable data than at those with reliable data (Supplementary **Fig. S5**). We analysed the response to WC for the inverse Simpson's concentration index and Pielou evenness in Nagoya Port. However, similar analyses were not performed in Matsunaga Bay because of the dearth of reliable data.

The trend of the inverse Simpson's concentration index with WC was not significant in the analysis for reliable data with combined specific and test communities in Nagoya Port, but decreased significantly in the analysis for reliable data with only test communities (**Table 3**), which excluded the six sites with unreliable data that had a low index at high WC (**Fig. 5c**). Analyses including both reliable and unreliable data also showed significant negative responses to WC (**Table 3**). Pielou evenness showed no significant response in the sets of reliable data. However, in the analysis including unreliable data with test communities N6 and N17, which plotted at low Pielou evenness and high WC (**Fig. 5d**), Pielou evenness showed a questionable negative response to WC ($t = -2.3, P = 0.044$) (**Table 3**).

Discussion

Our analyses of marine invertebrate communities at a regional scale and at two local sites revealed that taxonomic density (species density) was a sensitive index for assessing the quality of marine sediment. However, although the inverse Simpson's concentration index and Pielou evenness were shown to respond to sediment variability in the regional dataset, they could be insensitive or respond falsely when a low number of individuals was observed, and when more than one community co-existed at a local site. These results from two local sites should serve as a warning when using diversity indices, although these indices can provide a good understanding of how communities respond to sediment degradation. We need to understand how these indices collapse with an insufficient number of individuals observed or by co-existing communities, and we need better strategies for the ecological assessment of sediment quality using diversity indices at the local scale in marine areas.

The analyses for the regional dataset show that WC had the greatest impact on the taxonomic density of benthic invertebrate communities compared to other variables, although grain size and organic matter content are also thought to affect the richness of benthic invertebrates^{e.g. 13,14}. The high contribution of WC likely reflects its physical effects on sediment related to its structure. The optimal range of WC for the burrowing activity of benthic invertebrates is between the densest (hardest; around 25% WC) and loosest (softest; around 40% WC) possible packing of sediment^{15,16}. A WC exceeding the upper limit of this optimal range may indicate sediment that is too soft for the burrowing activity of benthic invertebrates, resulting in the negative effects of WC. The relatively higher impact of WC in the low-frequency group than in the other groups (**Table 2**) means that taxonomic density of this group declines more than the other groups through a reduction in richness by the degradation of sediment quality.

The relatively high standard deviation of random effects against the effect size of WC in GLMMs means that the importance of unmeasured variables on taxonomic density increased. Although salinity is a possible factor affecting the diversity of benthic invertebrates¹⁷, it was not measured in the regional dataset. However, our results showed a relatively high standard deviation of random effects on the taxonomic density in high-frequency groups. This may be a response to salinity in this group.

The analyses for the regional dataset also showed that an increase in WC caused a decrease in the inverse Simpson's concentration index and Pielou evenness in reliable data. Our results are consistent with previously identified responses to sediment degradation³. Because the low values in these indices occurred in communities with a few dominant species, these results suggest a scenario in which the benthic community would be dominated by a few species in soft sediment.

We found a response in species density to WC at two local sites without any dependence on data type. However, in Nagoya Port, there was a more moderate response in the analysis for the river-mouth and test communities combined than for the test community alone (**Fig. 5b**). This moderate response may result from low species density at low WC at the Fujimae intertidal flat. In addition, it was important to isolate the test community from other communities to find a statistically significant response in the inverse Simpson's concentration index at the port (**Table 3**). This would be caused by the low inverse Simpson's concentration index at N8 and the Fujimae intertidal flat (**Fig. 5c**). The influence from these sites is judged to occur as a result of masking by co-existing variable communities. Conversely, we found a questionable significant negative response to WC in the analyses for data type including unreliable data in the inverse Simpson's concentration index and Pielou evenness (**Table 3**), which may be caused by the unreliable (i.e. based on small sample sizes) low values at high WC (**Fig. 5c, d**). Masking by co-existing variable communities, and unreliable responses resulting from small sample size, obviously influenced the statistical results for species density, the inverse Simpson's concentration index, and Pielou evenness.

The influence of small sample size on the inverse Simpson's concentration index can be explained by a theoretical framework⁶. The index, which reflects the number of dominant species¹¹, is predicted to decline or remain unchanged in response to low species density when taxonomic density has a sensitive negative response (**Fig. 6a**). However, the index can be underestimated when there is a small sample size (**Fig. 6b**). Therefore, the response we observed is not a real response reflecting community structure but a false response that negatively correlates with sediment degradation, as can be seen in the questionable results from Nagoya Port. This influence of small sample size may also affect the Pielou evenness at this port. However, low Pielou evenness in unreliable data is a rare case (**Appendix S2**). Pielou evenness tended to be high, approaching 1.0, in the unreliable data of the regional dataset (Supplementary **Fig. S2**) and Matsunaga Bay (Supplementary **Fig. S4**). This bias can be explained as a possible result of small sample size. This should serve as a warning that false or insensitive responses in the inverse Simpson's concentration index and Pielou evenness may occur if data reliability is not checked.

The influence of small sample size shows the importance of assessing data reliability when calculating these indices. Sample coverage is an index that standardises the number of taxa represented by the individuals observed by completeness¹⁸. The sample coverage in reliable data was close to 1.0 (complete) and greater than that in the unreliable data in the three datasets used in this study. Although the cut-off in this study between reliable and unreliable data was not based on any scientific evidence, the value of 50 individuals was useful for judging the reliability of the estimates of these indices and whether they accurately reflect community structure. In addition, sample coverage was useful for plotting the degree of accuracy in two dimensional figures. Although the rarefaction curve is a more direct way to show the estimation accuracy for the inverse Simpson's concentration index, sample coverage has advantages in its usefulness for judging data reliability and for plotting in a figure.

Finding and excluding co-existing communities are important when analysing real responses to sediment degradation in a community. However, because the spatial variability in community structure is dependent on location, it is not easy to predict before field sampling. In Matsunaga Bay, the river-mouth community on the intertidal flat at the mouth of a small river was separated out by the dominance of the polychaete *Simplisetia erythraeensis*, which can be found in brackish environments (WoRMS: <http://www.marinespecies.org/>). It is easy to predict that this intertidal flat would have distinct communities in this bay from its location near a small freshwater river, as confirmed during field sampling. However, in Nagoya Port, although we expected the benthic community on the Fujimae intertidal flat to have a specific structure because of its location before the analysis, N8 and N20 were not predictable because of the complex spatial pattern of benthic communities. This complexity included different community types at N6 and N17, and at sites where no individuals were sampled. This complexity might be caused by the specific environmental characteristics of an urbanised marine area¹⁰ and might be difficult to determine in this area before field sampling. Multivariate analysis is required to distinguish between a test community and other communities. The use of multiple distances is recommended for finding how communities differ, both in species composition and in the proportions of individuals in each taxon.

In conclusion, we showed that species density responded sensitively to sediment degradation; however, the responses of indices of biological community structure, i.e. inverse Simpson's concentration index and Pielou evenness, were sometimes misleading because of inaccurate estimates with small sample sizes, and were insensitive at the local scale because of masking by variable

communities. Because indices for community structure provide a good understanding of how communities respond to sediment degradation, which cannot be provided by species density, ecological approaches using these indices have merits for assessing sediment quality because they are more realistic under field conditions² and because they reduce uncertainties^{19,20}. The possible misleading and insensitive results must be avoided to prevent diluting these merits. We recommend that these indices be used in local assessments if it is possible to obtain a sufficient sample size, and if co-existing communities can be differentiated.

Materials And Methods

Sample collection.

Surface sediment for biological and chemical analysis was obtained by using Ekman–Birge and Smith–McIntyre bottom samplers at sample locations (**Fig. 1**). Sampling depths for these samplers were about 20 cm²¹. Benthic invertebrates were obtained by sieving the sediment sample through 1-mm mesh in the field. The invertebrates were preserved in 10% formalin and returned to the laboratory. Biologists in the laboratory identified the taxa sampled and counted the numbers of individuals. Individuals were identified to the lowest taxonomic unit possible and corrected to the accepted names in the World Register of Marine Species (WoRMS: <http://www.marinespecies.org/>). Because the bottom sampler used and the number of samplings differed among observations within datasets, the data to be used were selected before analysis (see **Appendix S1**).

Sediment for chemical analysis was obtained separately from the biological samples. The sediment was chilled on ice immediately and transported to the laboratory. All three datasets include sediment temperatures. Oxidation–reduction potential (ORP) was also measured in Matsunaga Bay by using an ORP sensor (IM-32P; DKK-TOA Co. Ltd., Tokyo, Japan). Salinity in the water above the sediment surface was measured in Nagoya Port by using multiple-parameter water-quality meters (AAQ; Alec Electronics Co. Ltd., Hyogo, Japan). Chemical analyses of sediments were performed in the laboratory. Water content (WC, %), median sediment particle size (D_{50} , mm), total organic carbon (TOC, g kg⁻¹ dry wt.), and the molar carbon to nitrogen ratio (C/N) were measured in all three datasets (see **Appendix S3** for the details of chemical analysis).

Definitions and calculations of diversity indices.

Family was used as the lowest taxonomic level in the regional dataset. Species was used for local assessment in Matsunaga Bay and Nagoya Port (see **Appendix S3** for the definitions and calculations of target indices). The inverse Simpson's concentration index with the number of individuals observed, n , was calculated using the 'iNEXT' package²² in R (<https://www.r-project.org/>). Pielou evenness was calculated by dividing the Shannon index, output by the package, by log-transformed taxonomic density. We expected only rough values for the inverse Simpson's concentration index and Pielou evenness when the number of individuals observed was small. To assess this possible uncertainty, we distinguished between data for 50 or more individuals (reliable data), and data for fewer than 50 individuals (unreliable data). In addition, sample coverage, an index of sample completeness¹⁸, was also calculated using the 'iNEXT' package to assess whether sample size was sufficient.

Analysis for potential responses of diversity indices.

We analysed the family density, inverse Simpson's concentration index, and Pielou evenness for all data in the regional dataset. In addition, we applied the inverse Simpson's concentration index and Pielou evenness to only reliable data because of the possibility that these indices were less accurate when they included unreliable data. We also analysed the densities for three groups of families categorised as low- (present at ≤ 6 sites out of the total of 65 sites observed), intermediate- (present at between 7 and 25 sites), and high-frequency groups (≥ 26 sites) to find which groups were most affected by sediment quality. These divisions between groups were selected so that the mean numbers of families in observations for each group would be relatively even.

We used a generalised linear model with a random intercept (generalised linear mixed model, GLMM) for these family densities. We applied a Poisson error with a log function for all families and for the low-frequency group. We applied a binomial error with a logistic function for the intermediate- and high-frequency groups. We used a linear model with a Gaussian distribution for the analyses of the inverse Simpson's concentration index and Pielou evenness. The explanatory variables considered were those

related to sediment quality along with several other variables. The sediment-quality variables were WC, D_{50} , and TOC, which have been considered in previous studies^{12–14}, the interaction between TOC and the C/N ratio, which relates to sediment contaminants²³, and sediment temperature. Water depth and latitude were also considered because of their known effects on the diversity of marine benthic invertebrates^{24–27}. WC and D_{50} were log-transformed before analysis because of their wide ranges.

Areal sampler size was considered as a categorical variable with two possible values: large (i.e. the Smith–McIntyre bottom sampler) or small (the Ekman–Birge bottom sampler). Although the overall family density can be adjusted for sampler size by a rarefaction technique, the family density for the three frequency groups could not be adjusted because the rarefaction technique uses the overall relative abundance of individuals observed. Therefore, to consistently analyse the family density for the overall data and for each frequency group, we treated the areal sampler size as a sampling effect. This effect was included not only for taxonomic density but also for the inverse Simpson's concentration index and Pielou evenness, to test how their accuracies were affected by the sampler size.

We used a technique that averages the candidate models to visualise plausible explanatory variables among many possible good models^{28,29}. Before the averaging, we assessed the collinearity and goodness-of-fit of each model (see **Appendix S3** for the preparation of candidate models and model averaging). The average estimate of parameters for explanatory variables and their unconditional standard error were calculated by using the Akaike weight^{28,29}. Although we show the 95% confidence intervals for the effects of explanatory variables, we discuss them not only in terms of their significance but also in terms of how the possible explanatory variables change between model sets. The maximum likelihoods and standard deviations of regression coefficients were estimated using the 'glm' function for GLMs in R for the linear model, and the R package 'glmmML' for the GLMMs³⁰.

Assessment at two local sites.

The assessments of two local datasets employed the sediment variable that had the most impact in the regional dataset. Before determining the responses, we tested for differences in community structure in each area. After identifying a community for testing the response, we compared how the response of diversity indices changed for the other communities.

At each local site we first performed two cluster analyses: Sørensen dissimilarity was used to understand how species compositions differed, and Euclidean distance after a $\log(n + 1)$ transformation was used to understand how community structures differed in terms of abundance. We applied an unweighted pair-grouping method using arithmetic averages for the clustering based on Sørensen dissimilarity, and Ward's method for Euclidean distance. Then, we used distance-based redundancy analysis (dbRDA) to understand the relationship between community structure and environmental variables. Water depth and five sediment variables (WC, D_{50} , TOC, C/N, and sediment temperature) were used as common explanatory variables for Matsunaga Bay and Nagoya Port. Sample size and ORP were added for Matsunaga Bay. Salinity was added for Nagoya Port. The Euclidean distance after $\log(n + 1)$ transformation was used to define community patterns. An initial analysis was performed using forward stepwise tests of all environmental variables to choose the environmental variables that might best explain the patterns³¹. Goodness-of-fit was examined using the Akaike information criterion (AIC). Models with variance inflation factors of 10 or greater were rejected. We performed dbRDA on the variables selected for the most parsimonious model. The contributions of variables in the selected model were tested by permutational multivariate analysis of variance (PERMANOVA³²). These analyses revealed the communities that had different species compositions or specific structures in minor environments (i.e. "minor communities"). Cluster analysis was performed using the 'hclust' function in R, and PERMANOVA and dbRDA were performed using the 'adonis' and 'rda' functions in the R package 'vegan'³³, respectively. PERMANOVA was performed with 9999 permutations.

A community for testing the response to sediment variables was selected as a major community through two cluster analyses and dbRDA in each area. We analysed the trends of species density, the inverse Simpson's concentration index, and Pielou evenness against the sediment variable that had the most impact in the regional dataset, for the overall data and the reliable data, and for the data excluding minor communities, to test for the effects of unreliable data and minor communities on the diversity indices. We fit a generalised linear model using a Poisson error with a log function for species density, and normal error with a linear function for the inverse Simpson's concentration index and Pielou evenness. Species density at two sampling sites, M10 and M15, in Matsunaga Bay was adjusted to the number of individuals in an area corresponding to the size of a Smith–McIntyre

bottom sampler (see Supplementary **Table S5**) using a rarefaction technique⁶. In Nagoya Port, because no individual was obtained at a sampling site, N5, where different samplers were used, no adjustment was necessary.

Declarations

Acknowledgements

The authors thank Drs. Graeme Batley and Stuart Simpson from the Commonwealth Scientific and Industrial Research Organization (CSIRO), Australia, for their valuable comments, and Shota Okura of the Port and Airport Research Institute, Japan, for his discussions regarding the data analysis.

Author Contributions

S.H., K.M., and A.C. conceived the ideas and designed methodology; S.H., R.N., and Y.N. collected the data; S.H. analysed the data; S.H., K.M., and A.C. led the writing of the manuscript. All authors contributed critically to the drafts.

References

1. Chariton, A. A., Pettigrove, V. & Baird, D. J. Ecological assessment. in *Sediment quality assessment: a practical guide* (eds. Simpson, S. L. & Batley, G. E.) 346 (CSIRO publishing, 2016).
2. Chapman, P. M. Integrating toxicology and ecology: Putting the 'eco' into ecotoxicology. *Mar. Pollut. Bull.***44**, 7–15 (2002).
3. Johnston, E. L. & Roberts, D. A. Contaminants reduce the richness and evenness of marine communities: A review and meta-analysis. *Environ. Pollut.***157**, 1745–1752 (2009).
4. Gotelli, N. J. & Colwell, R. K. Quantifying biodiversity: Procedures and pitfalls in the measurement and comparison of species richness. *Ecol. Lett.***4**, 379–391 (2001).
5. Smith, W. & Grassle, F. Sampling properties of a family of diversity measures. *Biometrics***33**, 283–292 (1977).
6. Chao, A. *et al.* Rarefaction and extrapolation with Hill numbers: A framework for sampling and estimation in species diversity studies. *Ecol. Monogr.***84**, 45–67 (2014).
7. Gray, J. S., Elliott, M. & edition, S. *Ecology of Marine Sediments*. (Oxford University Press Inc, 2009).
8. Momota, K. & Nakaoka, M. Influence of different types of sessile epibionts on the community structure of mobile invertebrates in an eelgrass bed. *PeerJ***2017**, e2952; 10.7717/peerj.2952 (2017).
9. Momota, K. & Nakaoka, M. Seasonal change in spatial variability of eelgrass epifaunal community in relation to gradients of abiotic and biotic factors. *Mar. Ecol.***39**, e12522; 10.1111/maec.12522 (2018).
10. Todd, P. A. *et al.* Towards an urban marine ecology: characterizing the drivers, patterns and processes of marine ecosystems in coastal cities. *Oikos***128**, 1215–1242 (2019).
11. Hill, M. O. Diversity and evenness: a unifying notation and its consequences. *Ecology***54**, 427–432 (1973).
12. Hosokawa, S. Inference of sediment characteristics relating to the species richness of benthic marine animals, by considering regional and local variabilities. *J. Japan Soc. Civ. Eng. Ser. B***270**, 65–78 (2014). in Japanese with English abstract
13. Pearson, T. H. & Rosenberg, R. Macrobenthic succession in relation to organic enrichment and pollution of the marine environment. *Oceanogr. Mar. Biol. Annu. Rev.***16**, 229–311 (1978).
14. Ellingsen, K. E. Soft-sediment benthic biodiversity on the continental shelf in relation to environmental variability. *Mar. Ecol. Prog. Ser.***232**, 15–27 (2002).
15. Sassa, S. & Watabe, Y. Threshold, optimum and critical geoenvironmental conditions for burrowing activity of sand bubbler crab, *Scopimera globosa*. *Mar. Ecol. Prog. Ser.***354**, 191–199 (2008).
16. Sassa, S., Watabe, Y., Yang, S. & Kuwae, T. Ecological geotechnics: Role of waterfront geoenvironment as habitats in the activities of crabs, bivalves, and birds for biodiversity restoration. *Soils Found.***53**, 246–258 (2013).
17. Chapman, P. M. & Wang, F. Assessing sediment contamination in estuaries. *Environ. Toxicol. Chem.***20**, 3–22 (2001).

18. Chao, A. & Jost, L. Coverage-based rarefaction and extrapolation: Standardizing samples by completeness rather than size. *Ecology***93**, 2533–2547 (2012).
19. Batley, G. E., Burton, G. A., Chapman, P. M. & Forbes, V. E. Uncertainties in sediment quality weight-of-evidence (WOE) assessments. *Hum. Ecol. Risk Assess.***8**, 1517–1547 (2002).
20. Batley, G. E. & Simpson, S. L. Introduction. in *Sediment quality assessment: a practical guide* (eds. Simpson, S. L. & Batley, G. E.) 346 (CSIRO publishing, 2016).
21. U.S. EPA. *Methods for collection, storage and manipulation of sediments for chemical and toxicological analyses: Technical manual*. EPA 823-B-01-002 (2001).
22. Hsieh, T. C., Ma, K. H. & Chao, A. iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods Ecol. Evol.***7**, 1451–1456 (2016).
23. Hosokawa, S., Naito, R. & Nakamura, Y. Spatial patterns of concentrations of Cu, Zn, Cd, and Pb in marine sediments from Japanese port areas. *Reg. Stud. Mar. Sci.***35**, 101135; 10.1016/j.rsma.2020.101135 (2020).
24. Hillebrand, H. Strength, slope and variability of marine latitudinal gradients. *Mar. Ecol. Prog. Ser.***273**, 251–267 (2004).
25. Gray, J. S. Species richness of marine soft sediments. *Mar. Ecol. Prog. Ser.***244**, 285–297 (2002).
26. Rex, M. A. *et al.* Global-scale latitudinal patterns of species diversity in the deep-sea benthos. *Nature***365**, 636–639 (1993).
27. Witman, J. D., Etter, R. J. & Smith, F. The relationship between regional and local species diversity in marine benthic communities: A global perspective. *Proc. Natl. Acad. Sci. U. S. A.***101**, 15664–15669 (2004).
28. Anderson, D. R., Burnham, K. P. & Thompson, W. L. Null hypothesis testing: Problems, prevalence, and an alternative. *J. Wildl. Manage.***64**, 912–923 (2000).
29. Burnham, K. P. & Anderson, D. R. *Model selection and multimodel inference: a practical information-theoretic approach*. (Springer, 2002).
30. Broström, G. & Holmberg, H. Generalized linear models with clustered data: Fixed and random effects models. *Comput. Stat. Data Anal.***55**, 3123–3134 (2011).
31. Oksanen, J. Multivariate analysis of ecological communities in R: vegan tutorial. <https://www.mooreecology.com/uploads/2/4/2/1/24213970/vegantutor.pdf> (2015).
32. Anderson, M. J. A new method for non-parametric multivariate analysis of variance. *Austral Ecol.***26**, 32–46 (2001).
33. Oksanen, J. *et al.* Community Ecology Package 2.5-7. <https://cran.r-project.org/web/packages/vegan/vegan.pdf> (2020).

Tables

Table 1. Composition of benthic invertebrate communities as determined from three datasets and from the data for the low- (≤ 6 observations per 65 sites), intermediate- (7–25 observations), and high-frequency groups (≥ 26 observations) of families in the regional dataset. The values are the numbers of families in the regional dataset and numbers of species in the Matsunaga Bay and Nagoya Port datasets.

	Number of families or species recorded					
	Polychaeta	Malacostraca	Bivalvia	Gastropoda	Others	Total
Regional	38	32	20	14	39	143
Low	19	30	15	13	36	113
Intermediate	11	2	4	1	3	21
High	8	0	1	0	0	9
Matsunaga Bay	39	24	14	14	15	106
Nagoya Port	56	15	26	11	14	122

Table 2. Summary of the regional dataset and results of average generalised linear mixed models (GLMMs). Shown are the effect sizes of log-transformed water content (WC), the standard deviation of random effects (σ), and their ratio, for GLMMs in the overall dataset, and in the low-, intermediate-, and high-frequency groups. σ is significant in the overall dataset and all groups. WC is significant in the overall dataset and low frequency group but not significant in other two groups.

	No. of families in observation			Effect size in GLMMs		
	Mean \pm S.D.	Min	Max	WC	σ	WC/ σ
				Mean (95% CI)	Mean (95% CI)	
Overall	12.4 \pm 8.2	0	36	-0.34 (-0.58 to -0.10)	0.62 (0.46 to 0.78)	-0.55
Low	3.8 \pm 3.9	0	16	-0.54 (-0.92 to -0.16)	0.67 (0.45 to 0.88)	-0.81
Intermediate	3.9 \pm 3.2	0	15	-0.29 (-0.74 to 0.15)	0.87 (0.61 to 1.14)	-0.33
High	4.7 \pm 2.5	0	9	-0.23 (-0.70 to 0.24)	1.18 (0.84 to 1.51)	-0.20

Table 3. Summary of generalised linear model results for two local datasets: Matsunaga Bay and Nagoya Port. Shown are the responses of diversity indices to sediment water content. N.S. means that the response was not significant. Results in parentheses are questionable because they result from unreliable data (fewer than 50 individuals sampled).

Dataset	Data type analysed		Species density		Inverse Simpson's conc.		Pielou evenness	
	Data reliability	Community	<i>N</i>	Response	<i>N</i>	Response	<i>N</i>	Response
Matsunaga Bay	Reliable & unreliable data	A river-mouth & test communities	30	Negative	30	(N.S.)	30	(N.S.)
	Unreliable data	Test community	25	Negative	25	(N.S.)	25	(N.S.)
Nagoya Port	Reliable & unreliable data	Specific & test communities	22	Negative	22	(Negative)	18	(N.S.)
	Reliable & unreliable data	Test community	17	Negative	17	(Negative)	13	(Negative)
	Reliable data	Specific & test communities	16	Negative	16	N.S.	16	N.S.
	Reliable data	Test community	11	Negative	11	Negative	11	N.S.