










Seasonality of primary production explains the richness of pioneering benthic communities

Received: 30 November 2023

Accepted: 18 September 2024

Published online: 27 September 2024

 Check for updates

Matteo Cecchetto ¹✉, Agnès Dettai², Cyril Gallut ³, Matthias Obst⁴, Piotr Kuklinski⁵, Piotr Balazy ⁵, Maciej Chelchowski ⁵, Magdalena Małachowicz⁵, Anita Poćwierz-Kotus⁵, Małgorzata Zbawicka⁵, Henning Reiss⁶, Marc P. Eléaume^{2,3}, Gentile Francesco Ficetola ⁷, Christina Pavloudi ⁸, Katrina Exter ⁹, Diego Fontaneto ^{10,11} & Stefano Schiaparelli ^{1,12}

A pattern of increasing species richness from the poles to the equator is frequently observed in many animal taxa. Ecological limits, determined by the abiotic conditions and biotic interactions within an environment, are one of the major factors influencing the geographical distribution of species diversity. Energy availability is often considered a crucial limiting factor, with temperature and productivity serving as empirical measures. However, these measures may not fully explain the observed species richness, particularly in marine ecosystems. Here, through a global comparative approach and standardised methodologies, such as Autonomous Reef Monitoring Structures (ARMS) and DNA metabarcoding, we show that the seasonality of primary production explains sessile animal richness comparatively or better than surface temperature or primary productivity alone. A Hierarchical Generalised Additive Model (HGAM) is validated, after a model selection procedure, and the prediction error is compared, following a cross-validation approach, with HGAMs including environmental variables commonly used to explain animal richness. Moreover, the linear effect of production magnitude on species richness becomes apparent only when considered jointly with seasonality, and, by identifying world coastal areas characterized by extreme values of both, we postulate that this effect may result in a positive relationship in environments with lower seasonality.

The relationship between global environmental variables and the structure of benthic communities has been a key focus of ecological research in the last century, with the formulation of multiple hypotheses to explain the perceived latitudinal gradient of biodiversity¹. The endeavour of finding an explanation to ‘Latitudinal Diversity Gradients’ (LDG) has a long history² and builds on the observation of a general increase in species diversity from the poles to the equator³. However,

this effect was often found to be not monotonic, with diversity not continuously decreasing with latitude, studded with exceptions based on the taxonomic group considered⁴ and the peculiarities of specific regions (e.g. Antarctica⁵). Attempts to identify the cause of such patterns detected different ecological variables¹, showing conflicting results depending on both the chosen measures of biodiversity, and the candidate ecological drivers themselves⁶.

A full list of affiliations appears at the end of the paper. ✉ e-mail: matteocecchetto@gmail.com

These inconsistencies probably derive, to some extent, from the fact that the observed diversity might be shaped by a variety of factors (e.g. biotic interactions, evolutionary influences, historical processes, etc.), regardless of the environmental settings, making it even more complicated to study². The intrinsic influence of evolutionary events, operating on a larger temporal and spatial scale, affects species distributions⁷. On top of this, the structure of the survey design and the allocation of sampling frames, i.e., where samples are collected, contribute significantly to the final outcomes of any investigation performed⁸. The application of traditional methodologies, such as those relying on morphological identifications, can be affected by low precision and reproducibility⁹, and this could further obfuscate the mechanisms underlying these processes.

Here, we evaluate the influence of environmental variables on shaping the richness of pioneering metazoan communities settling on artificial structures at the global scale. This is performed, for the first time, on a DNA metabarcoding dataset of shallow hard-bottom sessile metazoan communities recovered from Autonomous Reef Monitoring Structures (ARMS) deployed globally, assembled from publicly available and published datasets, including the ARMS Marine Biodiversity Observation Network programme¹⁰ (ARMS-MBON), as well as additional structures deployed in the Southern Ocean. These communities are strongly related to local environmental features compared to 'older' or '*climax*' assemblages¹¹. We hypothesise that the number of species recovered through globally distributed ARMS is determined by the amount and seasonality of resources, allowing the support of more species.

Of the many hypotheses formulated to explain the LDG, this one pertains to the 'Ecological limits' category, as defined by Pontarp et al.². The ecological limit can be defined as the 'limit to the number of individuals and/or taxa that can coexist within an ecosystem due to abiotic settings and biotic interactions such as competition for limited resources'². However, identifying the specific mechanisms that determine the relationship between the ecological constraints and the diversity measure is a complicated endeavour. Uncovering these mechanisms require a proper formulation of the theoretical framework, the identification of suitable empirical measures of the explanatory environmental variables (i.e. what exactly is the 'limit' and how it is measured) and a correct testing of the proposed predictions¹².

In addition to the geographic area or the long-term climatic stability, the energy availability is often mentioned as a main ecological limit to the number of species that can coexist in an environment^{2,6}. Temperature and productivity are the variables that historically received more attention as representative of this limit⁶. However, the relationship between temperature and diversity is complex. This includes, for example, its effect on physiological processes¹³, ultimately influencing the higher speciation rate in the tropics. Moreover, temperature itself can influence the total productivity present in an environment, especially in the terrestrial realm where the amount of resources is mostly driven by temperature and water¹⁴. Nonetheless, the effect that temperature has on defining the geographic distribution of species often resulted in a positive correlation, reflecting the higher richness found in the warm waters of the tropics. On the other hand, productivity is often represented in terms of magnitude (e.g. annual means of net primary productivity), which, especially in marine environments, can show an inverse relationship with diversity^{15,16}, bewildering our understanding of the relationship between productivity and diversity.

One aspect of primary productivity in marine ecosystems that has rarely been considered in describing the observed diversity of benthic metazoans refers to its seasonality¹⁷. The seasonality of oceanic primary productivity captures the regularity of production in the first layers of the ocean and has important implications for nutrient and carbon cycling¹⁸, influencing the capacity of marine ecosystems to

uptake atmospheric CO₂^{19,20}. Different oceanic regions present patterns of more or less intense seasonality for a variety of reasons: at higher latitudes for the reduction (or even absence) of primary production throughout the winter¹⁸ or in coastal waters for the seasonal input of riverine and aeolian nutrients or the presence of upwelling²¹. These factors can disrupt the recognised co-variation between seasonality and sea surface temperature, leading to considerable differences in seasonality patterns between ocean basins at lower latitudes²¹. Thus, the presence of a marked seasonality in the primary production of coastal waters can be due to the interplay of a variety of environmental conditions, as well as to the specific geomorphological and hydrological characteristics of the study site, independently from the main physicochemical characteristics of specific areas (such as temperature and salinity). Although this aspect of productivity has been studied for decades²², less attention has been dedicated to uncovering specific indices that can represent the effect seasonality has on benthic communities^{17,23}.

In this study, we test the role of the interaction between the seasonality of primary productivity and its magnitude, as a suitable ecological limit driving the richness of settling benthic metazoans on a global scale. We assume that the interplay between the aforementioned variables acts as a proxy to energy availability, shaping the resource availability to larvae and propagules. A Hierarchical Generalised Additive Model (HGAM), including separate smoothers for both seasonality and magnitude of net primary production, is validated, showing a strong negative relationship between seasonality and benthic richness. Prediction errors, calculated following a cross-validation approach, are lower for models including seasonality, indicating that the seasonal oscillation of resources has the strongest effect in this relationship. Moreover, coastline areas of the world exhibiting higher resource availability overlap with regions of known high biodiversity, matching the distribution of tropical coral reefs. By adopting standardised and reproducible methodologies, such as those obtained using DNA metabarcoding techniques^{10,11}, more robust data can be produced in order to predict the structure and diversity of pioneering hard-bottom communities under evolving environmental conditions. The application of ARMS at the global level, following a systematic approach and in a clear setting, provides a tool for understanding the influence that environmental variables have on the development of future assemblages.

Results

Out of the total 146 samples included in the original dataset, 140, corresponding to 116 ARMS, passed the quality control steps of the bioinformatic analyses. Specifically, the ARMS from the Aegean Sea were removed before proceeding with the rest of the analyses. The remaining samples were recovered from 14 different ecoregions²⁴, with a varying number of site and sampling replicates (the number of ARMS, Table 1).

The bioinformatic analyses performed on the entire dataset yielded a total of 1,065,469 denoised sequences and 6069 molecular Operational Taxonomic Units (mOTUs). The highest mOTU richness was observed in samples deployed within tropical latitudinal zones, whereas the lowest was observed in polar regions as well as in some temperate regions (Fig. 1). Among the environmental variables, the Normalised Seasonality Index (NSI) and sea surface temperature mean (SST mean) exhibited the highest Pearson's correlation coefficients with mOTU richness, while the SST range and total Net Primary Productivity (total NPP) showed the lowest coefficients (Supplementary Fig. 1).

Some environmental variables, including depth and months of deployment, showed high levels of collinearity (Supplementary Fig. 2). The highest correlation was observed between productivity measures (e.g. NPP mean and CHL mean), and between NSI and SST mean, but other variables, including SSS mean and SST range, were correlated

Table 1 | ARMS deployment design for each ecoregion

Ecoregion	Time range (years)	# Months of deployment	# Sites/year	# ARMS per site	# ARMS	Depth (metres)	Dataset origin
North and East Barents Sea	2018–2020	12–13	2	1	2	14	ARMS-MBON
Northern Norway and Finnmark	2019–2020	18	1	2	2	2	ARMS-MBON
North Sea	2018–2020	12–14	4	2–3	10	4–25	ARMS-MBON
Baltic Sea	2013–2015	26	2	2–3	5	8	Pearman et al. ²⁷
Celtic Seas	2019–2020	12	1	3	3	12	ARMS-MBON
South European Atlantic Shelf	2013–2014	12–13	3	2–3	8	11–12	Pearman et al. ²⁷
Western Mediterranean	2013–2014	18	3	3	9	17	Pearman et al. ²⁷
Adriatic Sea	2014–2015	13	2	3	6	9	Pearman et al. ²⁷
Black Sea	2013–2014	13	3	2–3	8	7–8	Pearman et al. ²⁷
Northern and Central Red Sea	2013–2020	14–24	10	1–6	34	10	Pearman et al. ²⁷ ARMS-MBON Carvalho et al. ²⁹
Southern Red Sea	2015–2017	24–33	5	3	15	8–10	Carvalho et al. ²⁹
Hawaii	2016–2018	23	1	6	6	4	Nichols et al. ⁵¹
East Antarctic Wilkes Land	2014–2017	12–24–36	4	1–2	6	20	This Paper
Ross Sea	2015–2018	12–24–36	3	2	6	25	This Paper

This table reports the most significant differences in the study design of each ecoregion, following Spalding et al.²⁴ ecoregions' naming, with rows ordered by latitude of the centre of the ecoregion. The field '#sites/year' shows the number of sites per year of deployment investigated in each region, and corresponds to the grouping factor adopted as a random effect in the HGAM model. Sites where multiple structures have been deployed in different years are grouped separately. Multiple values in the field 'depth' represent the range of depth at which the structures were deployed. The 'dataset origin' field refers to the publication the corresponding samples have been first presented. More information on the specifics of each study are reported in the published papers or the Supplementary Methods and Supplementary Data 5.

with several other variables included in the dataset (Supplementary Fig. 2).

Generalised additive model

Prior to model validation, the presence of spatial autocorrelation was assessed by fitting two HGAM models with separate smoothers for NSI and total NPP, and including or not the grouping factor 'site/year' as a random effect. The presence of the grouping factor as a random effect resulted in a significant reduction of the spatial autocorrelation, a substantial decrease of the Akaike Information Criterion (AIC) value and no discernible patterns in the residuals compared to the HGAM model without the random effect (Supplementary Fig. 3).

Out of all variables, only SSS range, months and depth of deployment showed Variance Inflation Factor (VIF) values lower than 3, together with NSI and total NPP, and could all be included as fixed terms of a HGAM model with the aforementioned random effect (Table 2).

All the variables included in the model were statistically significant at the 5% level. Considering the fixed effects, the highest chi-square were observed for NSI and months of deployment (Table 2). The limited significance of the remaining variables can also be deduced by inspecting the confidence intervals of the smoothers, closely reaching the 0 partial-effect line (Supplementary Fig. 4). The number of mOTUs tended to increase with the months of deployment, but the pattern resulted not linear, and the increase was particularly strong after two years of deployment (Supplementary Fig. 4).

This model was compared with a simpler HGAM model fitted using only the energy availability variables (NSI and total NPP). This model showed a good fit and AIC values slightly higher than the former model (Table 3). Considering the limited significance of most non-collinear variables, and the non-linearity of the smoother for months of deployment, the lower AIC values could have been obtained through over-fitting, which is supported by the very similar values of percentage of explained deviance (Table 3). For this reason, the following analyses focus on the simpler model, only considering the variables included in the main hypothesis, NSI and total NPP, as proxies of energy availability. Validation of this model showed no discernible

pattern in the residuals, indicating normality of the same and a linear relationship between the fitted and observed response variables. The spread of the fitted values increased at higher values of the response variable (Supplementary Fig. 5).

The prediction of the combined effects of NSI and total NPP on the response variable revealed a predominantly negative effect of NSI (Fig. 2), which is also supported by the lower *p* values retrieved for both HGAM models (Table 2). This indicates a decreasing number of mOTUs with an increase in the seasonality of NPP, irrespective of the total NPP (Fig. 2). A negative effect was also observed with increasing total NPP. This effect was observed also in the initial model with all non-collinear variables. Nonetheless, the highest richness in the model was predicted for lower values of both seasonality and magnitude of production, reflecting the negative relationship between both smoothers and the number of mOTUs (Fig. 2).

Alternative models and assessment of the prediction error

The significance of NSI and total NPP was also assessed by comparing the prediction error of the aforementioned model with other HGAM models fitted with environmental variables commonly used to describe the richness of benthic communities.

Out of all these models, those including NSI and SST mean exhibited the strongest, linear relationships with mOTU richness, while the one including total NPP resulted in a complex non-linear and overall negative relationship, with the prediction values peaking at intermediate values of NPP (Fig. 3).

The prediction error, assessed by calculating the Root Mean Square Error (RMSE) values during the cross-validation approach, varied significantly among models (Fig. 4). The prediction error was higher for models including mean CHL and NPP (Fig. 3), and lower for SST mean, NSI and the energy availability model. The lowest RMSE values were obtained for the model including both NSI and total NPP. When predicting with fixed seasonality values on the validation datasets, the energy availability model exhibited a linear negative prediction of total NPP on the number of mOTUs, consistent with varying values of seasonality (Fig. 5). However, the effect was reduced at higher

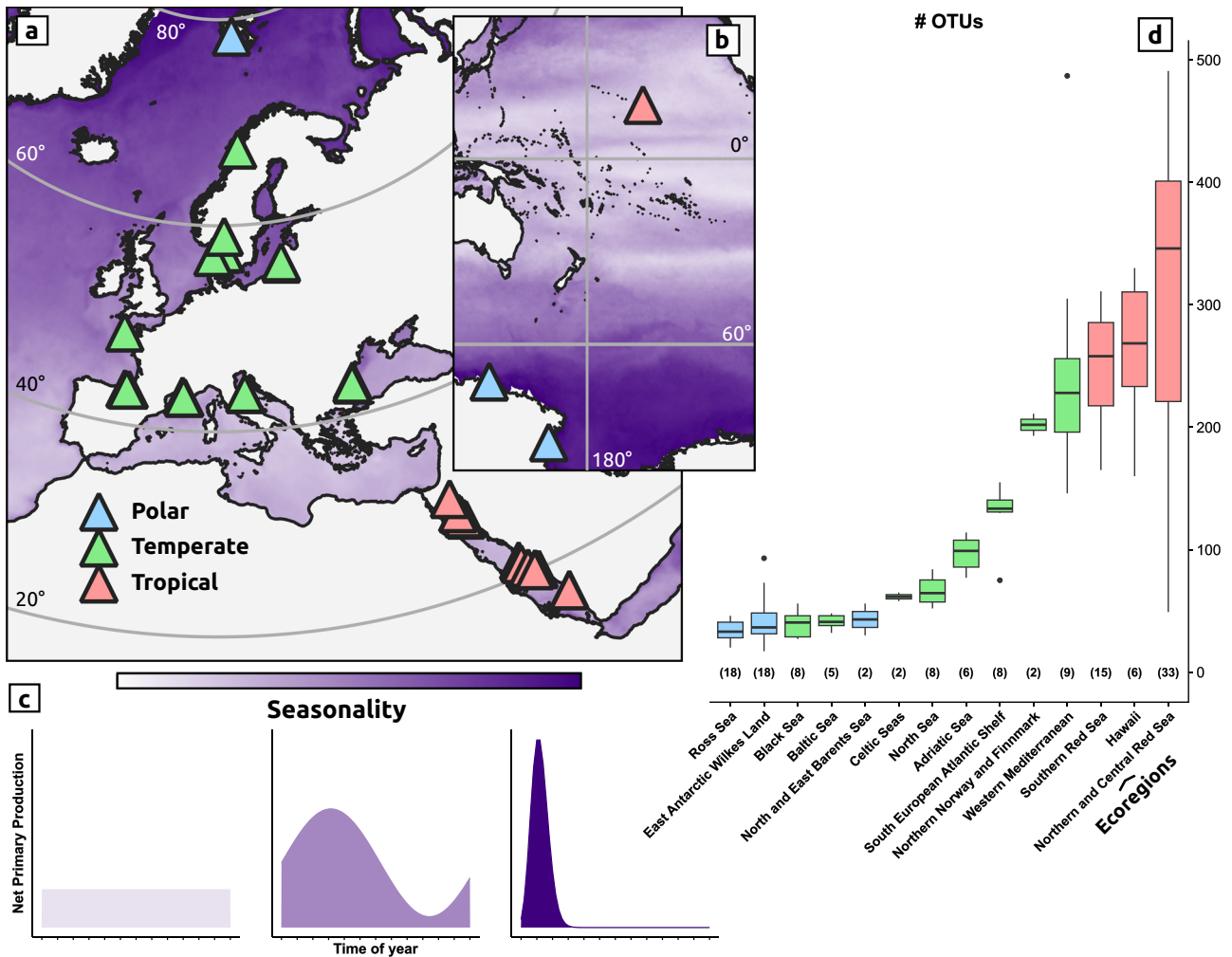


Fig. 1 | Deployment locations and richness of the ARMS investigated in this study. **a** Deployment locations for the Autonomous Reef Monitoring Structures (ARMS) investigated by Carvalho et al.²⁹ and Pearman et al.²⁷, as well as data obtained from the ARMS–Marine Biodiversity Observation Network (ARMS-MBON) programme¹⁰. **b** Deployment locations for the ARMS investigated by Nichols et al.³¹ as well as the ones from Antarctica. Purple tone shows the degree of seasonality, as described in box (c). **c** Schematic representation of the dynamics of Net Primary Production at different seasonality levels, following the same representation used in Berger and Wefer²². Lighter purple tones reflect a condition where primary production is constant throughout the year, whereas darker tones indicate higher seasonality, with most of the yearly production mostly occurring in a specific season. **d** Tukey-style boxplots showing the distribution of the number of molecular Operational Taxonomic Units (mOTUs) for the structures in each ecoregion investigated. The sample number onto which boxplots are calculated is shown in

brackets above the x axis labels, and refers to technical replicates. The lower and upper hinges correspond to the 25th and 75th percentiles, while the central line refers to the median. The upper whisker extends to the largest value no further than 1.5 * inter-quantile range (IQR) from the 75th percentile, while the lower extends to the smallest value at most 1.5 * IQR of the 25th percentile. All data points outside these ranges are plotted individually. These boxplots are shown only for the purpose of indicating the variability of mOTU richness across all the ecoregions. The triangles' and boxplots colours refer to the latitudinal zone in which the deployment took place, and the boxplots are ordered by the median of the number of mOTUs of the samples for each ecoregion. Ecoregions and latitudinal zones were defined following Spalding et al.²⁴. The map was created using QGIS⁷³ (version 3.14), the coastline contour was downloaded from the Global Self-consistent, Hierarchical, High-resolution Geography database⁷⁴ (GSHHG, version 2.3.7).

seasonality values, indicating a limited decrease of mOTU richness at higher values of seasonality (Fig. 5).

Identification of coastal regions with high energy availability

The procedure adopted to uncover areas of the world's coastline with high energy availability (low seasonality and high magnitude), produced 286,843 random points in the coastline buffer area. The vast majority of these points were in areas with low magnitude of production, relative to the dataset's extreme values of NSI and total NPP, and either low or high seasonality (-0.35 and -0.9 of NSI), reflecting the higher extent of these type of regions in the world oceans' surface (Supplementary Figs. 6–9). Areas with low seasonality (NSI < 0.4) and high magnitude (total NPP > 75 g m⁻² year⁻¹) are under-represented in this dataset, resulting in 5% extrapolation

areas in the energy availability model not supported by data (Fig. 6a). Only 9504 points were generated in raster pixels for these combinations of NSI and total NPP values in the world's coastlines, reflecting this condition (Fig. 6b).

The percentage of high energy availability points for each realm²⁴, indicates that temperate Australasia, eastern tropical coastline of Africa and South America and the Indo-Pacific include more coastline areas with these characteristics, followed by other temperate regions (Table 4). However, when examining the percentage at the ecoregion level, the pattern is more fragmented and many ecoregions of western and central Indo-Pacific, as well as the tropical Atlantic, show high cover of high energy availability areas (Fig. 6c). This is compared to the percentage of points, in the same ecoregions, with low seasonality and no filter for magnitude of production (Fig. 6d). The complete list of

ecoregions with the respective number of total and filtered points can be inspected in Supplementary Data 1.

Discussion

One of the main goals in ecology is to understand which environmental factors shape the observed species richness, and a clear knowledge of the mechanisms driving these phenomena is pivotal. By focusing on specific taxonomic groups, and using less reproducible and standar-

dised methodologies, the results, and thus the inferences coming from their interpretation, may lack the necessary setting for a comprehensive comparison across a wide range of habitats, regions, and taxonomic groups.

Since the proposal of ARMS as a quantitative sampling methodology for hard bottom pioneering artificial communities²⁵, a multitude of projects have been conducted employing these structures in different areas of the world. However, only recently, different organisations and research networks have worked together to plan and conduct simultaneous monitoring activities at a continental and global level^{10,26}. The first regional study on colonisation of ARMS, conducted at a continental scale, was published by Pearman et al.²⁷, whose samples were included in the analyses performed here. Pearman et al.²⁷ didn't find any clear latitudinal trend in diversity, but highlighted correlations with environmental variables, chosen due to their previously known influence on the distribution of marine species, their broad characterisation of sea surface waters and their availability under a standardised format (i.e. from satellite observations).

Here, the spatial scale was expanded at the global level, and more environmental variables were included, providing a more complete ecological setting. This allowed us to test whether the number of benthic species settling on artificial structures is determined by the availability of resources, rather than temperature or primary productivity alone, by integrating an aspect of the phenology of planktonic communities that has rarely been considered in this context, i.e. seasonality. The rationale behind this hypothesis builds on the concept that a marked seasonality in coastal waters can determine intense phytoplanktonic blooms during specific seasons, with a reduced temporal range of primary production throughout most of the year (in contrast to less seasonal regions, like in the tropics). The seasonal distribution of primary production ultimately influences the processing of organic matter that reaches deeper waters²⁸, regulating the availability of food to higher trophic levels. In highly seasonal environments with a reduced temporal range of primary production, there is likely to be higher competition due to time-limited availability of resources. Conversely, regions with lower seasonality, where primary production remains relatively constant throughout the year, could support the settlement and coexistence of additional species.

Table 2 | HGAM models variables' statistics

Smooth terms	HGAM all non-collinear			HGAM NSI and Total NPP		
	edf	chi-square	p value	edf	chi-square	p value
NSI	4.052	73.110	<2e-16	2.771	122.151	<2e-16
Total NPP	1.000	4.383	0.036324	1.001	4.753	0.029337
Months	5.117	21.194	0.000947	-	-	-
Depth	1.000	4.855	0.027565	-	-	-
SSS range	1.718	6.582	0.037561	-	-	-
Site (random effect)	24.714	105.486	<2e-16	34.816	245.086	<2e-16

Output of the summary.gam function of the R package mgcv⁷⁸. Estimated degrees of freedom (edf), the chi-square test and p values (two-sided chi-square test) refer to each smooth term in the fitted HGAM models. The 'site' variable represents the grouping factor used as random effect.

Table 3 | HGAM models' statistics

Models	edf	Deviance explained %	AIC
NSI + Total NPP	38.59	93.86	1389.11
NSI + Total NPP + All non-collinear	37.6	93.96	1386.63

Estimated degrees of freedom (edf), percentage of deviance explained and Akaike Information Criterion (AIC) for all models fitted in the model selection process.

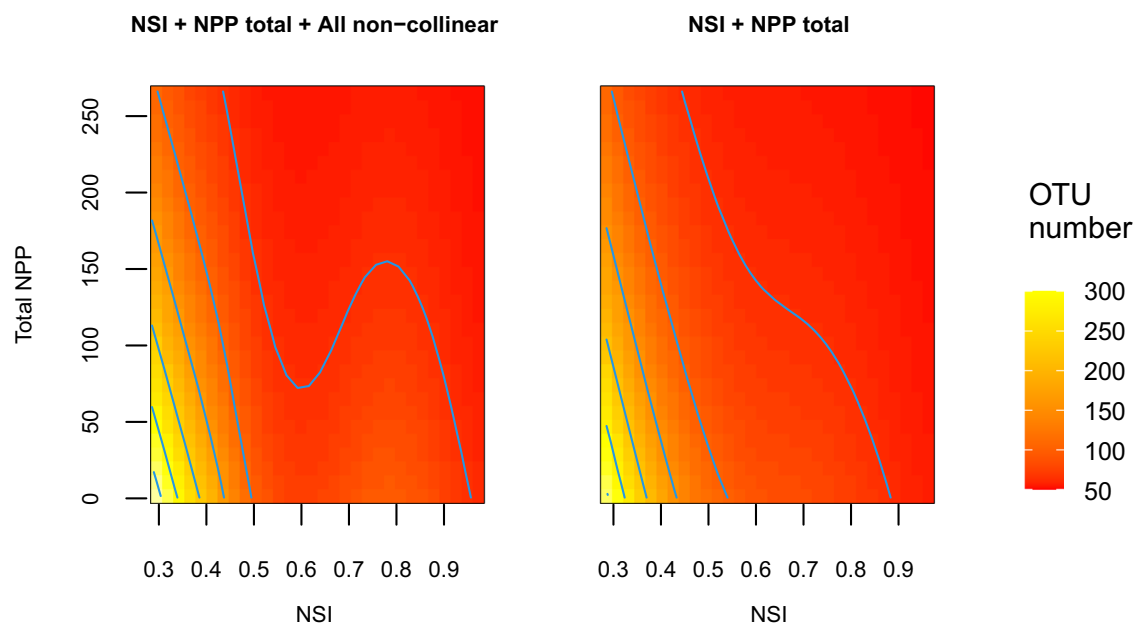


Fig. 2 | Contour plots of HGAM models predictions. The predictions of the number of mOTUs corresponding to the smoothers of total NPP and NSI in the HGAM models obtained. Yellow tiles indicate higher mOTUs number, red tiles

lower. The total NPP and NSI values are on the y and x axis respectively. The blue lines correspond to contour values of OTU number prediction, increasing by a factor of 50, as reported in the colour bar of the figure.

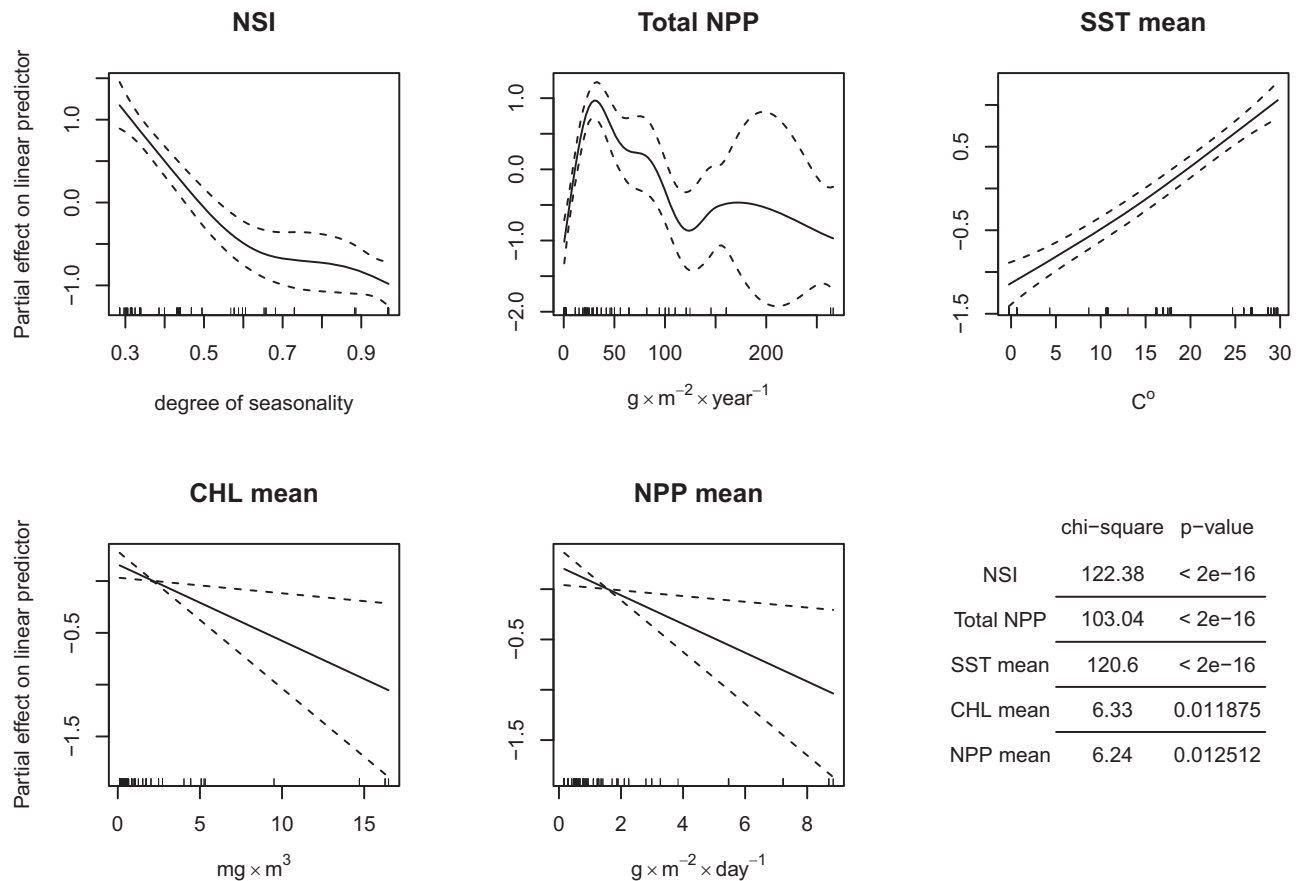


Fig. 3 | Partial effect plots of alternative HGAM models. Smoothers of the models fitted with single variables usually used in describing benthic metazoan richness (Sea Surface Temperatures, Chlorophyll and Net Primary Productivity means and

total), in addition to the NSI. The component smooth function is plotted on the scale of the linear predictor in the y-axis. The table reports the χ^2 and p values (two-sided chi-square test) of each single model.

In this context, the higher yearly magnitude of primary productivity would hardly provide any difference to the overall resource availability of benthic communities inhabiting regions with higher seasonality, as those resources would be already limited by the temporal dynamics of phytoplanktonic blooms. Differently, in less seasonal environments, a higher amount of primary productivity would have a greater impact as it would indicate a higher magnitude of resources distributed in a longer time frame, thus increasing, not only the annual quantity of productivity, but also the overall availability to benthic communities. This delineates four different ‘extreme’ scenarios characterised by either low or high seasonality combined with low or high productivity magnitude, where the regions showing low seasonality and high productivity could sustain the highest number of species.

However, our model detected a negative relationship between both NSI and total NPP and the number of benthic mOTUs (Fig. 2), with seasonality explaining most of the variation (Table 2). The highest number of species was predicted at low seasonality and magnitude of production, with a limited and negative effect of magnitude (Fig. 5). Although a positive relationship between richness and magnitude of production was here expected, at least for low seasonality levels, it is important to mention that the aforementioned outcome is not unknown in literature.

Other studies involving ARMS, but at a smaller spatial scale, uncovered a lower mOTU richness in the most productive areas, especially when similar measures of magnitude of production were taken into account. Pearman et al.²⁷ found a negative correlation between Shannon diversity and both the range of sea surface temperature and the mean concentration of chlorophyll. Although no clear linear relationship between SST range and mOTU richness was

detected here, similar outcomes were revealed regarding CHL mean (Fig. 3), reflecting the negative relationship between total NPP and mOTU richness in our HGAM models.

At an even smaller spatial scale, Carvalho et al.²⁹, using data from ARMS deployed along the Red Sea coastline, showed a negative correlation between both sea surface temperatures and particulate organic carbon, and benthic diversity of both sessile and motile metazoan communities. A significant decrease in benthic diversity was observed from the northern area of the Red Sea to the southern one, in accordance with the results of another paper³⁰, using the same structures but dealing with the microbial community, where a lower diversity was found in the southern region of the Red Sea. In fact, this region is characterised by higher concentrations of chlorophyll and is renowned for the higher seasonality of sea surface temperature and chlorophyll³¹, resembling the outcomes obtained in this paper, i.e. a negative relationship between magnitude of production and mOTU richness.

While consistent with the literature on ARMS studies, the negative relationship between total NPP and richness found here could also be explained by the absence of samples representing regions with low seasonality and high productivity. In general, regions at lower latitudes are under-represented in the dataset here analysed, compared to temperate ones, even though we have many replicated sites and samples in the Red Sea (Table 1 and Fig. 1), where extensive research using ARMS has been conducted^{27,29,30}. For this reason, higher values of magnitude of production are here only found in samples showing increasing seasonality, which, as mentioned earlier, would represent a condition that does not substantially increase the overall resource availability to benthic communities, due to the temporal limitation of

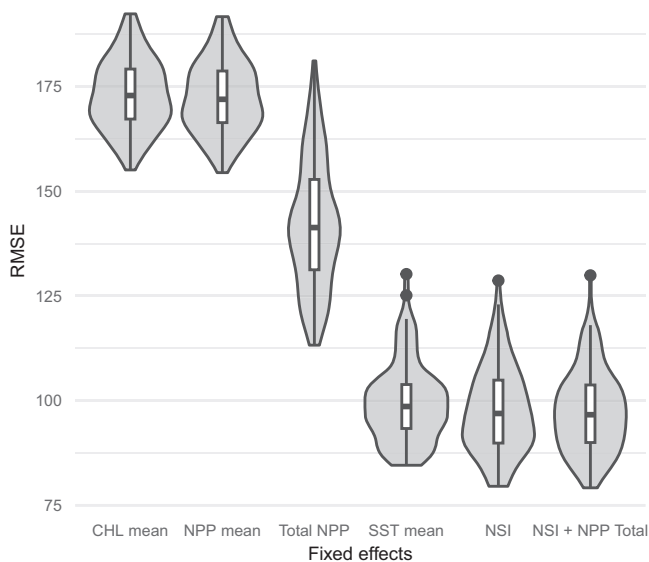


Fig. 4 | Violin plots of prediction errors. Violin plots showing the root-mean-square error (RMSE) values obtained in the cross-validation process for different HGAM models. Each violin refers to the RMSE values obtained for all 100 random validation datasets of a single model and has been ordered according to the median of all RMSE values. Tukey-style boxplots inside the violin plots show the 25th and 75th percentiles, as lower and upper hinges respectively, while the central line refers to the median. The upper whisker extends to the largest value no further than 1.5 * inter-quartile range (IQR) from the 75th percentile, while the lower extends to the smallest value at most 1.5 * IQR of the 25th percentile. All data points outside these ranges are plotted individually.

those resources. This is evidenced by the 5% extrapolation of the linear predictor of the model (Fig. 6a), indicating the presence of supporting data at varying degrees of total NPP only with increasing seasonality values (-0.5 to 0.6 of NSI), at which no significant increase in mOTU richness was predicted (Fig. 2). The identification of coastal regions including underrepresented environmental conditions (low seasonality and high magnitude of production) would improve our understanding of the effects of magnitude.

In fact, the ecoregions showing the highest percentage of low seasonality coastline areas are mostly located in the tropical and subtropical latitudinal zones, overlapping the global distribution of coral reefs³², showing a particularly high concentration in the Indo-Pacific, where the highest diversity of zooxanthellate Scleractinia can be found³³ (Fig. 6d). However, within this same distribution, only a limited number of ecoregions include high percentages of low seasonality and high magnitude of production areas (Fig. 6c), reaching a maximum of 12%, if we take into account the highest level of bioregionalisation, the realm (Table 4). These ecoregions are mostly located in the Northern coastline of Australia, overlapping both the eastern and western major coral reefs, the latter less studied, although they are comparable in extent to the former³⁴. Other areas include the Mozambique channel, where the highest diversity of the Indian Ocean reef-building corals can be observed^{33,35}, as well as the northern coastline of South America, either eastern or western, that, together with the Caribbean Sea, were found to host the highest biodiversity in Central to South America³⁶.

On the other hand, varying degrees of seasonality were thoroughly represented here. The highest seasonality of NPP were detected in polar regions, the Black and Baltic seas and Norway, with a distinct gradient from regions characterised by episodic (*sensu* Berger and Wefer²²) phytoplanktonic blooms (e.g. Antarctica and Svalbard), to areas with a more seasonal phenology (e.g. Black and Baltic seas), ending with regions characterised by oligotrophic waters and no

marked seasonal blooms (e.g. Red sea and Hawai'i). This trend also reflects the general assumptions of greater diversity in tropical regions in respect to both temperate and polar regions. In fact, lower richness was observed for the polar samples, followed by the Black Sea, in accordance with Pearman et al.²⁷, whereas the highest values were found in tropical regions, such as the Red Sea and the Hawai'i (Fig. 1d).

Prediction errors were found to be much higher in the HGAM model including measures of magnitude of production, such as total NPP only. On the other hand, the model with integrated NSI performed similarly to SST mean and NSI exclusively, with a much lower prediction error (Fig. 4). This condition is also reflected by the change of smoother for total NPP detected after its inclusion in an HGAM model together with NSI, resulting in lower degrees of freedom, highly differing from the smoother included in the HGAM model with total NPP only (Fig. 3). The linearity was found to remain consistent in the cross-validation approach (Fig. 5), suggesting that the primary role of the magnitude of NPP may be 'obscured' by the phenology of planktonic communities, at least for benthic communities directly relying on these resources, and may result effective only when integrated with its seasonality. However, the higher variation explained by seasonality in both HGAM models reinforce the idea that the temporal availability of food, instead of its magnitude, may represent the ecological limit for sessile metazoans, as seen in studies on deep-sea foraminifera¹⁷.

However, unexpected higher diversity was found for the samples from the Northern Norway and Finnmark ecoregion, highly resembling the mOTU richness found in the tropical ecoregions (Fig. 1d). This is particularly relevant, especially considering the high latitude at which those structures were deployed, and thus the high seasonality of primary productivity for this region. In fact, apart from the differences in deployment depth between these samples (i.e. from Norway) and the rest of the dataset (Table 1), the ARMS' location is characterised by high tidal currents exchanging water between two fjord systems with very different physical, chemical and hydrological conditions³⁷. These dynamic environmental conditions have already been associated with a higher diversity in macrobenthic species in respect to the neighbouring areas³⁸. The peculiar hydrological characteristics of this location complicate the relationship between the local primary production (of the specific area investigated) and benthic communities. This is due, in fact, to the higher sedimentation rates of organic matter and its frequent resuspension³⁸ present in the area, increasing the temporal availability of food resources for the developing communities.

This phenomenon is not unknown. Temperate and polar regions show moderate to high seasonal pulses of NPP, and thus consequently of particulate organic matter (POC). If the amount of POC produced during those pulses surpasses the consumers' capacity to feed upon it, it deposits in 'food banks', of more permanent and lower nutritional content³⁹. The resuspension of these food resources can sustain a more complex community, with more species of both depositional and suspension feeders. Such effects have already been observed, where suspension-feeding communities living in areas distant from the original location of this food source are supported by the availability of particulate matter suspended by currents⁴⁰. Thus, the relationship between food availability and diversity can be more complex if we take into account the peculiar dynamics of organic particle export from the surface to the seafloor⁴¹ of specific areas.

This also explains the inconsistencies observed in different studies conducted to investigate the relationship between the diversity of foraminifera communities and seasonality of net primary production. Gooday et al.⁴² detected a negative relationship between 'food availability' and diversity of deep sea benthic foraminifera communities, in contrast with the results observed in previous literature^{17,43}. Nematodes, foraminifera and molluscs richness positively correlates with latitude and POC fluxes to the seafloor in deep sea sites distributed globally, and not with surface primary productivity and POC production⁴⁴, suggesting that the

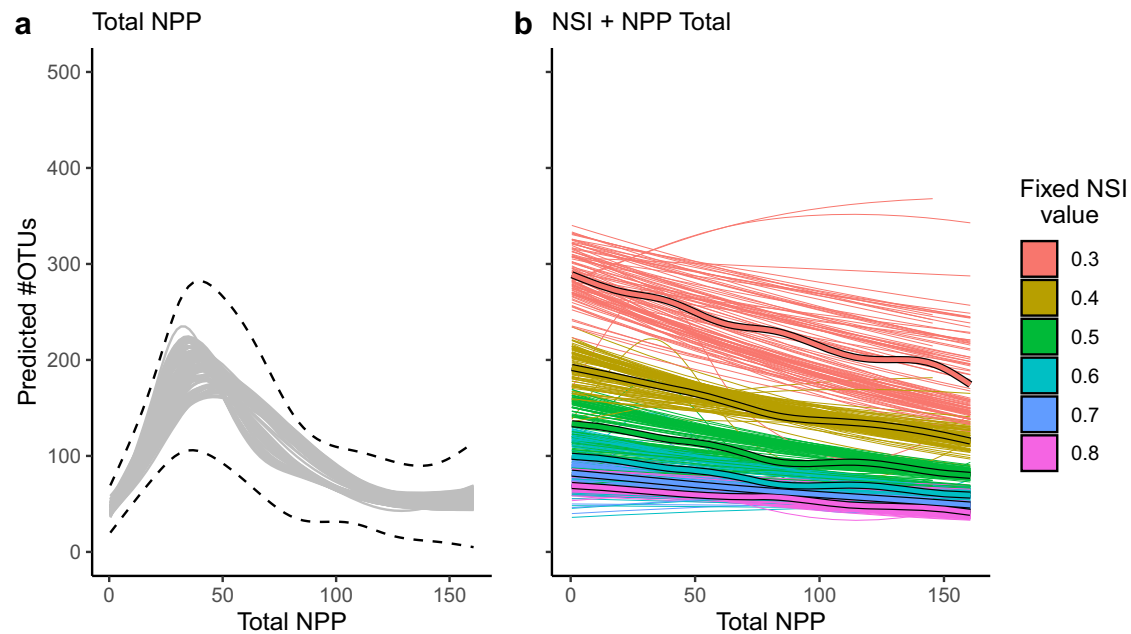


Fig. 5 | Smoother linearity of total NPP in HGAM models. Predictions effect (smoothers) for the models including the total NPP and tested on the validation datasets in the cross-validation process. **a** The prediction effects of the model including total NPP on the number of mOTUs. Dashed, black lines show the smoothing obtained using the command ‘geom_smooth’ in ggplot and the arguments ‘method = gam’ ‘formula = y - s(x)’ on the upper and lower confidence

intervals of each smoother in the validation datasets. **b** The prediction effect of the model including total NPP and NSI on the validation datasets with fixed values of the NSI at different levels of seasonality (from 0.4 to 0.8). Solid, stroked lines are produced using the command ‘geom_smooth’ in ggplot and adopting the arguments ‘method = gam’ and ‘formula = y - s(x)’. The unit for total NPP is reported in the ‘Methods’ section.

geographic mismatch, and renown decoupling⁴⁵, between the surface production and the effective export to the seafloor could explain, at least partly, these incongruences.

Another aspect to take into account refers to the relationship between mOTU richness and SST. SST mean was strongly related to mOTU richness, resulting in low prediction errors, comparable to the HGAM models including NSI exclusively or together with total NPP (Fig. 4). This outcome is not surprising, especially if we consider the high collinearity between SST mean and NSI, a well-known condition²¹, which has been detected also for this dataset (Supplementary Fig. 2). However, the high correlation observed here between these two variables may be due to the aforementioned under-representation of tropical regions. In fact, the correlation between temperature and seasonality of NPP is reduced at lower latitudes, thanks to the other parameters generally influencing the seasonality (e.g. upwelling or riverine inputs)²¹. By providing a more complete dataset, the collinearity between these two environmental variables could be reduced, resulting in a proper setting for a variable selection process, allowing to assess the significance of NSI over SST mean in explaining richness at the global level.

Overall, the application of ARMS appears to be well suited in studies investigating the relationship between benthic diversity and the environmental settings at all spatial scales. In fact, these structures allow the recovery of a variety of species belonging to a very specific community (i.e. metazoan pioneering artificial communities of shallow hard-bottom substrata). This community deeply relies on the environmental settings and dynamics of the local primary productivity, making them particularly useful as experimental subjects, especially in light of the mechanistic approach envisioned by Pontarp et al.². More precise and meaningful results could be obtained by providing a complex, multidisciplinary approach that gathers the logistic opportunities and expertise of multiple researchers, thus providing a monitoring programme that specifically focus on the effects of environmental parameters and biological dynamics on the pioneering assemblages of hard substrata throughout the world^{10,46}, including

multiple sites with contrasting or different environmental settings in each investigated region.

However, research programmes focusing on the development of this kind of studies have only recently been carried on at a global spatial scale and in a systematic way, taking advantage of the application of very specific and standardised methodologies, i.e. the ARMS, together with highly reproducible methodological techniques^{10,26}. Still, an important aspect to keep in mind when performing such programmes pertains to some characteristics of the study design, mainly the chosen depth of investigation and the operational deployment time range. Here, depth and months of deployment resulted statistically significant in a HGAM model that included different environmental variables, however, no significant interpretation of the effects on benthic communities can be discerned, except for a general increase of richness with months of deployment. Nonetheless, the influence of depth and months of deployment is surely present, but the geographical scale of the dataset may obscure it.

Another aspect to consider is the sample and site replication, here limited for some ecoregions in respect to others, due to the aggregation of data from studies conducted with different research objectives, designs and capabilities. Although no defined number of necessary replicates can be strictly determined, it is noteworthy to mention that a certain variability must be provided when modelling global distribution of benthic richness, and a lack of such variation could lead to biased interpretations. The higher than expected diversity recovered from the samples located in the ‘Northern Norway and Finnmark’ ecoregion could be an example, where the inclusion of multiple sites could unveil a different overall situation for the region, being characterised by a complex coastline and diversified environmental conditions, typical of fjord systems, that do not necessarily match the specific conditions here examined.

The representativeness of diverse regions, when considering the global spatial scale, is similarly important. Here, the diversity of temperate ecoregions included, in contrast to the limited number of tropical ecoregions, reduced the ability to investigate some aspects of

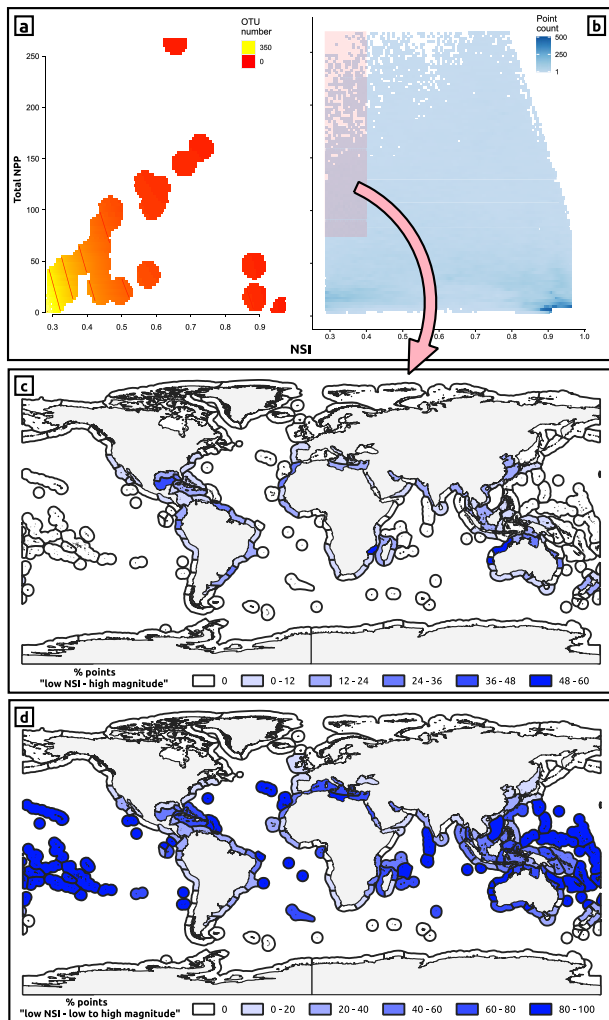


Fig. 6 | Graphic representation of the percentage of random point falling on coastal areas with high energy availability in each ecoregion. **a** 5% extrapolation areas for the mOTU richness prediction of the smoothers in the model including NSI and total NPP. **b** Density plot of all random points generated in the 0.04 degrees coastal areas of the Global Self-consistent, Hierarchical, High-resolution Geography database (GSHHG) world coastlines plotted with the corresponding total NPP (y axis) and NSI (x axis) values from the global rasters of those variables. Each tile can include a minimum of 1 point. Red shaded area represents the high energy availability combinations of total NPP and NSI values (NSI < 0.4 and total NPP > 75 g m⁻² year⁻¹), and matches the low confidence area of the prediction in box (a). **c** Ecoregions are coloured based on the percentage of number of points falling on high energy availability coastal areas in respect to the total number of random points. **d** As per box (c), but based on the percentage of points falling on low seasonality areas (NSI < 0.4), irrespective of total NPP values.

energy availability in the tropical latitudinal zone. However, additional samples from studies involving ARMS in tropical, or sub-tropical zones could not be included in this study due to the limited deployment time range²⁵, the adoption of different amplifying primers⁴⁷ or the absence of metabarcoding sequences⁴⁸. Future global comparison of ARMS samples will need to integrate more diverse sites in different locations, increasing the variability of regions from all latitudinal zones, while also selecting data with a higher site replication per region, provided the validity of the adopted bioregionalisation.

Acknowledging the influence that seasonality of primary production has on the diversity of hard-bottom, pioneering metazoan communities could improve our efforts to preserve biodiversity and our understanding of the impacts of anthropogenic pressure,

especially if we consider the different drivers of seasonality and the direct impacts humans have on them (e.g. eutrophication). At a larger spatial and temporal scale, this would help to model or infer the state of future assemblages' diversity under the influence of climatic changes, including the ongoing processes of atlantification⁴⁹ and topicalization in wide oceanic areas of the world, and/or to assess and monitor the effect of climatic oscillations on communities inhabiting key areas affected by these conditions⁵⁰. The adoption of ARMS is particularly useful in this context due to their flexible deploying potentiality, allowing to inspect the communities developing on these structures at different time ranges and thus, at different environmental conditions.

Methods

ARMS sequences

The ARMS sequences (140 samples, corresponding to 116 structures) analysed in this paper were gathered from both published resources, including Carvalho et al.²⁹, Pearman et al.²⁷ and Nichols et al.⁵¹, and from new data obtained in the context of different research programmes working at high latitudes in both Poles and the European continent, from the ARMS-MBON programme¹⁰ and this study. Only the samples obtained from structures deployed for at least one year were included, in order to analyse communities that had the chance to experience all the different physical and biological conditions of the water column in all different seasons. These sequences were downloaded from the respective data repositories cited in the aforementioned papers and incorporated in the pipeline with the new data presented here. More information on the deployment, retrieval, processing and DNA extraction, amplification and sequencing of the latter is presented in the Supplementary Methods. The same information for the samples obtained from the published papers can be obtained from the aforementioned studies^{27,29,51}.

Geographic extent of dataset

The ARMS samples from Carvalho et al.²⁹, Pearman et al.²⁷ and Nichols et al.⁵¹ were collected from 8 different ecoregions, all in the Northern Hemisphere and never exceeding 60 degrees of latitude (Table 1 and Fig. 1). The total 14 ecoregions and three latitudinal zones are defined following Spalding et al.²⁴. These ecoregions cover different seas and oceans, specifically the Red, Baltic, Black, North, Mediterranean and Barents seas, the Hawai'i islands and Southern Ocean, spanning from 78° north to 74° south of latitude (Fig. 1). Samples from the Southern Ocean were deployed near two research stations, namely 'Mario Zucchelli' station, in Terra Nova bay in the Victoria Land, and 'Dumont D'Urville' station in Terre Adélie.

The samples included in the analyses derive from structures deployed for at least one year at shallow depths (less than 25 metres). Due to the lack of any processed motile fraction in some of the samples' sources, only the sessile fraction was included. Furthermore, this allows investigating a more 'defined' community and reducing the amount of secondary consumers included in the analyses.

Bioinformatic analyses

Raw sequences from specific samples were processed differently from the rest of the dataset in the first steps of the pipeline due to structural differences of the format in which they were available: specifically, sequences from 'East Antarctic Wilkes land' (ENA accession numbers ERR12209601 to ERR12209606) were initially reverse complemented and subsequently demultiplexed using the vsearch⁵² and cutadapt⁵³ softwares, respectively, whereas for the samples from Carvalho et al.²⁹, as well as other samples from the ARMS-MBON programme, no primer removal was performed as the fastq files available were already trimmed from them. Following these pre-processing steps, all paired reads from the entire dataset were length truncated to 200 and 190 bp for the forward and reverse reads, respectively. Sequences from the two reads were merged with no ambiguous bases allowed and 2 maximum

Table 4 | High energy availability coastal areas percentage for each realm

Realms	# High energy availability points	# Total random points	% High energy availability
Temperate Australasia	865	7181	12.05
Tropical Atlantic	1866	15,664	11.91
Tropical Eastern Pacific	256	2700	9.48
Western Indo-Pacific	1403	18,354	7.64
Central Indo-Pacific	2727	40,367	6.76
Temperate Northern Pacific	1076	27,043	3.98
Temperate South America	301	12,607	2.39
Temperate Northern Atlantic	989	42,904	2.31
Temperate Southern Africa	21	1164	1.8
Arctic	0	89,340	0
Eastern Indo-Pacific	0	3609	0
Southern Ocean	0	25,910	0

Percentage of random points falling on high energy availability areas (low seasonality–high magnitude of production) for each realm, according to Spalding et al.²⁴.

differences in the alignment. The merged sequence length was restricted to 311–313 bp for the samples that were already available without primers, whereas the length restriction was set to those values plus the length of the primers (26 bp) and primer tags, if present. These steps were performed using *vsearch*, whereas primer removal was done using *cutadapt*, allowing one error maximum in the alignment for each primer. Samples with less than 200 sequences at this point were removed from the analyses. All samples were concatenated into one fastq file, sequences with a maximum expected error greater than 1 were removed, converted into fasta format and length filtered to 311–313 bp using *vsearch*, which was also used for the dereplication, and removal of chimeras, performed using the *uchime2*⁵⁴ software.

A list of accession numbers of mitochondrial genome sequences stored at the NCBI (Supplementary Data 2) was used to download the corresponding fasta files of the mitochondrial cytochrome oxidase I (COI) region from a variety of animal phyla, using a script written with different tools of the Entrez Direct software suite⁵⁵, in addition to *samtools*⁵⁶ and *seqkit*⁵⁷ softwares. These sequences were aligned using the software *MACSE*⁵⁸ (tool ‘AlignSequences’) together with the primers from Leray et al.⁵⁹ and Geller et al.⁶⁰, the same used to amplify all the samples used in this study. The software *Megax*⁶¹ was used to inspect the alignment and extract the smallest coding region including the ‘Leray fragment’ of the COI, and one additional codon at the third end (thus obtaining an alignment of 318 bp). This was done in order to provide some ‘freedom’ to the following alignment step, which was done aligning the dereplicated dataset to that coding region. This final alignment was performed using the tool ‘EnrichAlignment’ in *MACSE*, which aligns each input sequence singularly and discards those that do not satisfy particular conditions (in this case, no frameshifts, insertions and stop codons were allowed). In order to speed up the alignment, the fasta file with the dereplicated sequences was split in multiple files with an identical number of sequences each, and run in parallel using the software tool *GNU parallel*⁶². This alignment procedure allowed us to identify the exact codon position of each nucleotide, and calculate the entropy values of each codon position in the sequences of the dataset, which is used in the denoising procedure proposed by Antich et al.^{63,64}. As the abundance annotation provided by *vsearch* at the dereplication refers to the total dataset, and thus might not reflect the true abundance at the DNA sequencing procedure, the denoising is performed for each group of samples analysed separately at PCR and sequencing (e.g. different sequencing runs), after recalculating the abundance of each sequence in each group of samples, using mainly *vsearch* and *seqkit* commands. Details on the procedure are available within the scripts. After denoising, a minimum abundance threshold of 0.005%, in respect to the total abundance for each group of samples,

was set, and all Amplicon Sequence Variants (ASV) with abundance of less than 5 sequences were removed from the analyses.

The molecular Operational Taxonomic Unit (mOTU) table (or count/abundance table) was assembled using the *vsearch* tool ‘search_extract’ on the dataset after the length filtering step and with the sequences aligned by *MACSE* as input. Clustering was then performed on the dereplicated, aligned and denoised sequences using the software *swarm*^{65,66} (version 3) and a resolution set to 13 due to the high mutation rate of the COI⁶³. The mOTU table earlier produced was then ‘collapsed’ according to the clustering results, using an R script modified from the script ‘*owi_recount_swarm.R*’ at the GitHub ‘*metabarpark*’ repository (https://github.com/metabarpark/R_scripts_metabarpark), which sums count values from each sequence belonging to the same cluster as obtained by the *swarm* algorithm.

Environmental parameters

The Moderate Resolution Imaging Spectroradiometer (MODIS) ancillary data on Chlorophyll (CHL, in mg m³), Sea Surface Temperature (SST, in °C), Photosynthetic Available Radiation (PAR, in E m² day⁻¹) and Diffuse Attenuation coefficient (DA, at 490 nm in units m⁻¹) was downloaded from the Ocean Productivity website (Oregon State University, <http://sites.science.oregonstate.edu/ocean.productivity/>) at 8-day and 0.083 degrees of resolution (or 2160 × 4320 grid size, -9 km at equator), from 2013 to 2020. With the same spatial resolution, but at 1-day of frequency, satellite data on Sea Surface Salinity (SSS, in PSU) was downloaded from the ‘Global Ocean Physics Reanalysis’ product⁶⁷ of the E.U. Copernicus Marine Service Information. The downloaded, raw data for SST, CHL and PAR were used for the calculation of net primary productivity (NPP, g m⁻² day⁻¹), and, consequently, the Normalised Seasonality Index (NSI), as described in the following paragraph. Long-term averages of mean values were calculated for all environmental variables, whereas long-term averages of the range values were additionally calculated for SSS and SST, as indicators of physical extremes. These variables have been historically used as descriptors of the richness and diversity of biological communities inhabiting marine environments, as they properly describe the physico-chemical and biological conditions of these ecosystems. Moreover, SST and NPP are usually investigated as empirical measures used to describe the ‘energy availability’ of marine environments. The sum of all NPP values (hereafter referred to as ‘total NPP’ g m⁻² year⁻¹) in a year for each satellite data point was preferred to the mean of NPP as representation of the magnitude of primary productivity, due to frequent absence of satellite data in polar areas in the winter months, ultimately influencing the final mean values.

Multiple indices have been proposed to quantify the annual variability of NPP, by taking into account the overall annual standard deviation in primary production²¹, the initiation and duration of phytoplanktonic bloom events (as reported in Cole et al.⁶⁸), or the 'production half-time'²². More recently, a new index (NSI) was proposed by Brown et al.¹⁸ by normalising the seasonality index introduced by Berger and Wefer²² by half of the total number of possible observations in a year, allowing its comparison between regions with different magnitudes of primary production. Two scripts in shell (Ubuntu version 20.04.6 LTS) and R (version 4.2.2) were created for the download and pre-processing of satellite data, the calculation of the Vertically Generalised Production Model (VGPM) for oceanic NPP, following Behrenfeld and Falkowski⁶⁹, using an R version of the code provided by the Ocean Productivity website (<http://sites.science.oregonstate.edu/ocean.productivity/vgpm.code.php>), and the calculation of the Normalised Seasonality Index (NSI) following Brown et al.¹⁸, using GDAL⁷⁰ (Geospatial Data Abstraction Library) the R packages collection tidyverse⁷¹ and a modified version of the 'daylength' function from the geosphere package⁷². This was done to include satellite data for the Black Sea, which are lacking in the layers for the VGPM of oceanic NPP provided by the Ocean Productivity website. The obtained data was then further processed using QGIS⁷³ (version 3.14). All points falling inside the coastline contour of the Global Self-consistent, Hierarchical, High-resolution Geography database⁷⁴ (GSHHG, version 2.3.7) were removed, and a kriging interpolation was performed using SAGA CMD⁷⁵, with default parameters except for the variogram, for which the model '2.20114 + 2.75005 * x' was chosen, and the target size, which was set to 0.04 degrees, thus doubling the resolution of the satellite raw data. The interpolation was performed on 16 regions of the world, overlaying the samples' locations and extending on -5 degrees of both latitude and longitude around those locations.

All raster files for these layers were imported in QGIS and values extracted using the coordinates of each sampling location (Supplementary Data 3). In order to account for spatial autocorrelation and reduce issues due to pseudo-replication, a grouping factor was created by aggregating samples (and sites) based on a distance of 0.02 degrees as radius, thus with a diameter of 0.04 degrees that resembles the spatial resolution of the interpolated environmental layers. As some structures were deployed in the same geographical point, but at different years and time ranges, they were discriminated against in the grouping factor based on the year of deployment (and/or retrieval). This grouping factor will be later used in the statistical analyses, to account for spatial autocorrelation.

The following analyses were also conducted using satellite data corresponding to the exact time range of each ARMS' deployment and retrieval schedule, thus obtaining only the environmental conditions that each sample faced during its community development, without relying on the averages of a broader period. However, due to the absence of satellite data for the exact locations of the ARMS deployed in Antarctica, likely caused by the persistent presence of sea-ice in those periods, samples from the 'East Antarctic and Wilkes land' ecoregion could not be included, whereas satellite data from the 'Ross Sea' ecoregion mostly relied on interpolation of satellite points located far from the exact location of the structures. For this reason, the analyses were conducted on a broader time range (2013–2020), including the deployment-retrieval time ranges of all samples. The results from the same analyses, but conducted on the aforementioned exact time ranges' data, are included in the Supplementary Discussion and Supplementary Data 4.

Statistical analyses

A phyloseq⁷⁶ object was created in R using the mOTU table and the sample information file, including the environmental variables (Supplementary Data 5). Dot charts showing the ranges and frequency of values for each environmental variable were inspected to check for

outliers (Supplementary Fig. 10), however, no transformation was deemed necessary. The correlation between the number of mOTUs and each environmental variable was assessed using Pearson's correlation coefficient, and the distribution of the number of mOTUs in each ecoregion (following Spalding et al.²⁴) was inspected with boxplots.

The presence of spatial autocorrelation was assessed with the package ncf⁷⁷, by estimating the spline correlograms on the Pearson's residuals of two HGAM models fitted using the package mgcv⁷⁸, both including separate smoothers for NSI and total NPP, as explanatory variables of the number of mOTUs. The two models differ in the presence (or not) of the random effect used to account for the spatial dependency within each 'site' (i.e. the grouping factor applied at the same spatial resolution of the environmental layers), where multiple samples (ARMS) have been deployed. This random effect also differentiates structures that have been deployed at the same sites but at different time ranges, for example in cases of multiple deployments at different years. The effectiveness of the random effect to be used in the following model selection and validation was assessed by comparing the AIC of the two models.

In order to check whether other environment variables could be included in the model, the entire set of variables was compared in a correlation matrix, before performing a variable selection based on the Variance Inflation Factor (VIF) using the package olsrr⁷⁹. The variables with less than 3 VIF, including the NSI and total NPP, were then included in the model. The variable selection was performed also on the depth and number of months of deployment.

Model selection was conducted on a HGAM model fitted using a REstricted Maximum Likelihood (REML) estimation and assuming a Negative Binomial distribution with logarithmic link for the response variable. The grouping factor 'site' was included in the model as a random effect and all the non-collinear explanatory variables were included in the model using thin plate regression splines smoothers for each variable. Model selection was performed until all variables resulted statistically significant (i.e. <0.05). The final model was compared to the model that only includes NSI and total NPP. Contour plots of each model's predictions were produced and compared to assess the combined effect of NSI and total NPP, and the models were validated by comparing the output of mgcv's 'gam.check' function, e.g. histogram of residuals and plot of residuals vs. fitted values.

Prediction performance between the HGAM model including NSI and total NPP was compared with that of other HGAM models including single variables commonly used in literature to describe benthic metazoan richness such as SST, CHL, total and mean NPP, and NSI, following a cross-validation approach. For each ecoregion, 2 sites were randomly sampled and the corresponding samples used to create the training datasets. This was done to avoid over-representation of the samples from the Red Sea, which accounted for a high proportion of the dataset's ARMS (Table 1). Each training dataset was then used to fit the different HGAM models and the prediction was conducted on the remaining subset (the rest of the dataset not sampled previously, here called the validation dataset) using the 'predict.gam' function, from the mgcv package. The prediction was performed including the random effect, as the random selection of sites allows to preserve the entirety of samples in each selected site, and the errors of the prediction were assessed by calculating the RMSE, using the package Metrics⁸⁰, between the prediction values and those observed in the validation dataset. This was performed on 100 randomly generated subsets of the entire dataset. The list of site codes in each training dataset used for the analyses and figures here reported is available at Supplementary Data 6.

The linearity of total NPP at varying degrees of seasonality was assessed by modifying the NSI values of each generated validation dataset, producing 6 new datasets at different levels of the seasonality index (i.e. from 0.3 to 0.8). This allowed us to inspect the smoother of total NPP at different levels of seasonality (in the models including the

NSI) and compare it to the smoothers obtained by the model with only total NPP as explanatory variable.

Identification of high energy availability coastal areas

In order to identify which areas of the world oceans' surfaces included extreme values of both seasonality and productivity, for the environmental variables of total NPP and NSI, an interpolation spanning the entire world was performed after selecting which of 225 regions dividing the earth surface in equal areas overlapped the oceans, and performing an interpolation inside each of these regions. The same aforementioned SAGA CMD parameters were used, except for search range and radius, which were changed to 'local' and 0.65 degrees respectively, to reduce the interpolation time. All interpolated layers from all regions were merged into a single one using GDAL and imported in QGIS. For each combination of total NPP and NSI extremes, a polygon layer showing their distribution on oceanic surface waters was created in QGIS, adopting a series of tools that allowed us to filter specific raster values.

Concurrently, the identification of which ecoregions included the highest percentage of coastline areas with high 'energy availability' (low seasonality and high productivity magnitude), a different approach was undertaken. A buffer area of 0.04 degrees of span from the coastlines identified by the GSHHG high resolution database was created, and 10 thousand points maximum falling on this buffer area were randomly extracted for each ecoregion, with a minimum distance of 0.04 degrees between each other. The points falling on raster pixels with low (NSI < 0.4) and high (total NPP > 75 g m⁻² year⁻¹) values of NSI and total NPP, respectively, were extracted and counted. The identification of low and high values of seasonality and productivity was defined by the inspection of the dataset distribution of these values, thus corresponding to the highest and lowest values in the dataset here analysed. Finally, for each ecoregion, the percentage of number of points in high energy availability areas compared to the total number of points was calculated. The percentage of all points falling on low seasonality areas (NSI < 0.4), independently from the productivity values, was also calculated.

A more detailed description on the methods described in this section can be found in the Supplementary Methods.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The .fastq files of the additional samples included in the analyses (and corresponding to the samples deployed in the Southern Ocean) are available at the European Nucleotide Archive (ENA, <http://www.ebi.ac.uk/ena>)⁸¹ with the study accession number [PRJEB67891](https://doi.org/10.6084/m9.figshare.25957663). The raw ARMS-MBON sequence files used for this study are available from the ENA as well, with the study accession numbers [PRJEB37740](https://doi.org/10.6084/m9.figshare.25957663), [PRJEB33796](https://doi.org/10.6084/m9.figshare.25957663), [PRJEB37751](https://doi.org/10.6084/m9.figshare.25957663), [PRJEB37741](https://doi.org/10.6084/m9.figshare.25957663), [PRJEB48198](https://doi.org/10.6084/m9.figshare.25957663), [PRJEB48199](https://doi.org/10.6084/m9.figshare.25957663). The run accession numbers for all samples available at ENA are listed in the Supplementary Data 3, while all additional information for each sample (depth, environmental data, date, etc.) is available in Supplementary Data 5. All raster layers, corresponding to the satellite data downloaded, as well as the shapefiles created for the identification of high energy availability coastal areas are available at FigShare <https://doi.org/10.6084/m9.figshare.25957663>.

Code availability

All scripts and files used for the bioinformatic and statistical analyses, as well as those to produce the NSI raster layer, are available at the two GitHub repositories https://github.com/MatteoCe/ARMS_seasonal and https://github.com/MatteoCe/vgpm_nsi. The steps adopted to download

the .fastq files of the samples gathered from the published papers are described in the former repository.

References

1. Fine, P. V. A. Ecological and evolutionary drivers of geographic variation in species diversity. *Annu. Rev. Ecol. Evol. Syst.* **46**, 369–392 (2015).
2. Pontarp, M. et al. The latitudinal diversity gradient: novel understanding through mechanistic eco-evolutionary models. *Trends Ecol. Evol.* **34**, 211–223 (2019).
3. Chaudhary, C., Saeedi, H. & Costello, M. J. Marine species richness is bimodal with latitude: a reply to Fernandez and Marques. *Trends Ecol. Evol.* **32**, 234–237 (2017).
4. Fernandez, M. O. & Marques, A. C. Diversity of diversities: a response to Chaudhary, Saeedi, and Costello. *Trends Ecol. Evol.* **32**, 232–234 (2017).
5. Gili, J.-M., Coma, R., Orejas, C., López-González, P. J. & Zabala, M. Are Antarctic suspension-feeding communities different from those elsewhere in the world? *Polar Biol.* **24**, 473–485 (2001).
6. Storch, D., Bohdalková, E. & Okie, J. The more-individuals hypothesis revisited: the role of community abundance in species richness regulation and the productivity–diversity relationship. *Ecol. Lett.* **21**, 920–937 (2018).
7. 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. USA* **101**, 15664–15669 (2004).
8. Jeliakov, A. et al. Sampling and modelling rare species: conceptual guidelines for the neglected majority. *Glob. Change Biol.* **28**, 3754–3777 (2022).
9. Baird, D. J. & Hajibabaei, M. Biomonitoring 2.0: a new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. *Mol. Ecol.* **21**, 2039–2044 (2012).
10. Obst, M. et al. A Marine Biodiversity Observation Network for genetic monitoring of hard-bottom communities (ARMS-MBON). *Front. Mar. Sci.* **7**, 572680 (2020).
11. Curini-Galletti, M. et al. Patterns of diversity in soft-bodied meiofauna: dispersal ability and body size matter. *PLoS ONE* **7**, e33801 (2012).
12. Etienne, R. S. et al. A minimal model for the latitudinal diversity gradient suggests a dominant role for ecological limits. *Am. Nat.* **194**, E122–E133 (2019).
13. Peck, L. S. A cold limit to adaptation in the sea. *Trends Ecol. Evol.* **31**, 13–26 (2016).
14. Brown, J. H. Why are there so many species in the tropics? *J. Biogeogr.* **41**, 8–22 (2014).
15. Macpherson, E. Large-scale species–richness gradients in the Atlantic Ocean. *Proc. R. Soc. Lond. B Biol. Sci.* **269**, 1715–1720 (2002).
16. Witman, J. D., Cusson, M., Archambault, P., Pershing, A. J. & Mieszkowska, N. The relation between productivity and species diversity in temperate–arctic marine ecosystems. *Ecology* **89**, S66–S80 (2008).
17. Corliss, B. H., Brown, C. W., Sun, X. & Showers, W. J. Deep-sea benthic diversity linked to seasonality of pelagic productivity. *Deep Sea Res. Part Oceanogr. Res. Pap.* **56**, 835–841 (2009).
18. Brown, C. W., Schollaert Uz, S. & Corliss, B. H. Seasonality of oceanic primary production and its interannual variability from 1998 to 2007. *Deep Sea Res. Part Oceanogr. Res. Pap.* **90**, 166–175 (2014).
19. Hughes, L. Biological consequences of global warming: is the signal already apparent? *Trends Ecol. Evol.* **15**, 56–61 (2000).
20. Platt, T. & Sathyendranath, S. Ecological indicators for the pelagic zone of the ocean from remote sensing. *Remote Sens. Environ.* **112**, 3426–3436 (2008).

21. Lutz, M. J., Caldeira, K., Dunbar, R. B. & Behrenfeld, M. J. Seasonal rhythms of net primary production and particulate organic carbon flux to depth describe the efficiency of biological pump in the global ocean. *J. Geophys. Res. Oceans* **112**, C10011 (2007).
22. Berger, W. H. & Wefer, G. Export production: seasonality and intermittency, and paleoceanographic implications. *Glob. Planet. Change* **3**, 245–254 (1990).
23. Rex, M. A. et al. Global-scale latitudinal patterns of species diversity in the deep-sea benthos. *Nature* **365**, 636–639 (1993).
24. Spalding, M. D. et al. Marine ecoregions of the world: a bioregionalization of coastal and shelf areas. *BioScience* **57**, 573–583 (2007).
25. Leray, M. & Knowlton, N. DNA barcoding and metabarcoding of standardized samples reveal patterns of marine benthic diversity. *Proc. Natl Acad. Sci. USA* <https://doi.org/10.1073/pnas.1424997112> (2015).
26. Santi, I. et al. European marine omics biodiversity observation network: a strategic outline for the implementation of omics approaches in ocean observation. *Front. Mar. Sci.* **10**, 1118120 (2023).
27. Pearman, J. K. et al. Pan-regional marine benthic cryptobiome biodiversity patterns revealed by metabarcoding autonomous reef monitoring structures. *Mol. Ecol.* **29**, 4882–4897 (2020).
28. Francois, R., Honjo, S., Krishfield, R. & Manganini, S. Factors controlling the flux of organic carbon to the bathypelagic zone of the ocean. *Glob. Biogeochem. Cycles* **16**, 34–20 (2002).
29. Carvalho, S. et al. Beyond the visual: using metabarcoding to characterize the hidden reef cryptobiome. *Proc. R. Soc. B Biol. Sci.* **286**, 20182697 (2019).
30. Pearman, J. K. et al. Disentangling the complex microbial community of coral reefs using standardized Autonomous Reef Monitoring Structures (ARMS). *Mol. Ecol.* **28**, 3496–3507 (2019).
31. Raitsoo, D. E., Pradhan, Y., Brewin, R. J. W., Stenchikov, G. & Hoteit, I. Remote sensing the phytoplankton seasonal succession of the red sea. *PLoS ONE* **8**, e64909 (2013).
32. van Dam, J. W., Negri, A. P., Uthicke, S. & Mueller, J. F. Chemical pollution on coral reefs: Exposure and ecological effects. In *Ecological Impacts of Toxic Chemicals* (eds Sanchez-Bayo, F. et al.). Vol. 9, 187–211 (Bentham Science Publishers, Oak Park, 2011).
33. Veron, J., Stafford-Smith, M., DeVantier, L. & Turak, E. Overview of distribution patterns of zooxanthellate Scleractinia. *Front. Mar. Sci.* **1**, 81 (2015).
34. Gilmour, J. P. et al. The state of Western Australia’s coral reefs. *Coral Reefs* **38**, 651–667 (2019).
35. Obura, D. The diversity and biogeography of Western Indian Ocean reef-building corals. *PLoS ONE* **7**, e45013 (2012).
36. Miloslavich, P. et al. Marine biodiversity in the Atlantic and Pacific coasts of South America: knowledge and gaps. *PLoS ONE* **6**, e14631 (2011).
37. Eliassen, I. K., Heggelund, Y. & Haakstad, M. A numerical study of the circulation in Saltfjorden, Saltstraumen and Skjerstadfjorden. *Cont. Shelf Res.* **21**, 1669–1689 (2001).
38. Kokarev, V., Tachon, M., Austad, M., McGovern, M. & Reiss, H. Strong macrobenthic community differentiation among sub-Arctic deep fjords on small spatial scales. *Estuar. Coast. Shelf Sci.* **252**, 107271 (2021).
39. Campanyà-Llovet, N., Snelgrove, P. V. & Parrish, C. C. Rethinking the importance of food quality in marine benthic food webs. *Prog. Oceanogr.* **156**, 240–251 (2017).
40. Kaehler, S., Pakhomov, E., Kalin, R. & Davis, S. Trophic importance of kelp-derived suspended particulate matter in a through-flow sub-Antarctic system. *Mar. Ecol. Prog. Ser.* **316**, 17–22 (2006).
41. Jansen, J. et al. Abundance and richness of key Antarctic seafloor fauna correlates with modelled food availability. *Nat. Ecol. Evol.* **2**, 71–80 (2018).
42. Gooday, A. J., Bett, B. J., Jones, D. O. B. & Kitazato, H. The influence of productivity on abyssal foraminiferal biodiversity. *Mar. Biodivers.* **42**, 415–431 (2012).
43. Wollenburg, J. E. & Mackensen, A. Living benthic foraminifers from the central Arctic ocean: faunal composition, standing stock and diversity. *Mar. Micropaleontol.* **34**, 153–185 (1998).
44. Cordier, T. et al. Ecosystems monitoring powered by environmental genomics: a review of current strategies with an implementation roadmap. *Mol. Ecol.* **30**, 2937–2958 (2021).
45. Buesseler, K. O. The decoupling of production and particulate export in the surface ocean. *Glob. Biogeochem. Cycles* **12**, 297–310 (1998).
46. Convey, P. & Peck, L. S. Antarctic environmental change and biological responses. *Sci. Adv.* **5**, eaaz0888 (2019).
47. Ip, Y. C. A. et al. Seq’ and ARMS shall find: DNA (meta)barcoding of Autonomous Reef Monitoring Structures across the tree of life uncovers hidden cryptobiome of tropical urban coral reefs. *Mol. Ecol.* <https://doi.org/10.1111/mec.16568> (2022).
48. Couëdel, M. et al. New insights into the diversity of cryptobenthic *Cirripectes* blennies in the Mascarene Archipelago sampled using Autonomous Reef Monitoring Structures (ARMS). *Ecol. Evol.* **13**, e9850 (2023).
49. Górska, B., Gromisz, S., Legeżyńska, J., Soltwedel, T. & Włodarska-Kowalczyk, M. Macrobenthic diversity response to the atlantification of the Arctic Ocean (Fram Strait, 79° N)—a taxonomic and functional trait approach. *Ecol. Indic.* **144**, 109464 (2022).
50. Glynn, P. W., Maté, J. L., Baker, A. C. & Calderón, M. O. Coral bleaching and mortality in Panama and Ecuador during the 1997–1998 El Niño–Southern Oscillation event: spatial/temporal patterns and comparisons with the 1982–1983 event. *Bull. Mar. Sci.* **69**, 79–109 (2001).
51. Nichols, P. K., Timmers, M. & Marko, P. B. Hide ‘n seq: direct versus indirect metabarcoding of coral reef cryptic communities. *Environ. DNA* **4**, 93–107 (2022).
52. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4**, e2584 (2016).
53. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10–12 (2011).
54. Edgar, R. C. UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. *BioRxiv* <https://doi.org/10.1101/081257> (2016).
55. Kans, J. Entrez direct: E-utilities on the UNIX command line. In *Entrez Programming Utilities Help* (National Center for Biotechnology Information, 2024) https://www.ncbi.nlm.nih.gov/books/NBK25501/pdf/Bookshelf_NBK25501.
56. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
57. Shen, W., Le, S., Li, Y. & Hu, F. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS ONE* **11**, e0163962 (2016).
58. Ranwez, V., Harispe, S., Delsuc, F. & Douzery, E. J. P. MACSE: Multiple Alignment of Coding Sequences accounting for frameshifts and stop codons. *PLoS ONE* **6**, e22594 (2011).
59. Leray, M. et al. A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. *Front. Zool.* **10**, 1–14 (2013).
60. Geller, J., Meyer, C., Parker, M. & Hawk, H. Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. *Mol. Ecol. Resour.* **13**, 851–861 (2013).
61. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547 (2018).

62. Tange, O. GNU parallel—the Command-Line Power Tool. *Logix USENIX Mag.* **36**, 42–47 (2011).
63. Antich, A., Palacin, C., Wangensteen, O. S. & Turon, X. To denoise or to cluster, that is not the question: optimizing pipelines for COI metabarcoding and metaphylogeography. *BMC Bioinforma.* **22**, 177 (2021).
64. Antich, A., Palacin, C., Turon, X. & Wangensteen, O. S. DnoisE: distance denoising by entropy. An open-source parallelizable alternative for denoising sequence datasets. *PeerJ* **10**, e12758 (2022).
65. Mahé, F., Rognes, T., Quince, C., de Vargas, C. & Dunthorn, M. Swarm: robust and fast clustering method for amplicon-based studies. *PeerJ* **2**, e593 (2014).
66. Mahé, F., Rognes, T., Quince, C., de Vargas, C. & Dunthorn, M. Swarm v2: highly-scalable and high-resolution amplicon clustering. *PeerJ* **3**, e1420 (2015).
67. E.U. Copernicus Marine Service Information (CMEMS). Global Ocean Physics Reanalysis. Marine Data Store (MDS). <https://doi.org/10.48670/moi-00021> (2024).
68. Cole, H., Henson, S., Martin, A. & Yool, A. Mind the gap: the impact of missing data on the calculation of phytoplankton phenology metrics. *J. Geophys. Res. Oceans* **117**, 8030 (2012).
69. Behrenfeld, M. J. & Falkowski, P. G. Photosynthetic rates derived from satellite-based chlorophyll concentration. *Limnol. Oceanogr.* **42**, 1–20 (1997).
70. GDAL/OGR Contributors. *GDAL/OGR Geospatial Data Abstraction Software Library*. (Open Source Geospatial Foundation, 2024).
71. Wickham, H. et al. Welcome to the {tidyverse}. *J. Open Source Softw.* **4**, 1686 (2019).
72. Hijmans, R. J. *geosphere: Spherical Trigonometry*. (R Foundation for Statistical Computing, 2024).
73. QGIS Development Team. QGIS geographic information system. (Open Source Geospatial Foundation Project, 2024).
74. Wessel, P. & Smith, W. H. F. A global, self-consistent, hierarchical, high-resolution shoreline database. *J. Geophys. Res. Solid Earth* **101**, 8741–8743 (1996).
75. Conrad, O. et al. System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. <https://gmd.copernicus.org/preprints/8/2271/2015/gmdd-8-2271-2015.pdf>; <https://doi.org/10.5194/gmdd-8-2271-2015> (2015).
76. McMurdie, P. J. & Holmes, S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8**, e61217 (2013).
77. Bjornstad, O. N. *nfc: Spatial Covariance Functions* (R Foundation for Statistical Computing, 2024).
78. Wood, S. N., Pya, N. & Saeften, B. Smoothing parameter and model selection for general smooth models (with discussion). *J. Am. Stat. Assoc.* **111**, 1548–1575 (2016).
79. Hebbali, A. *olsrr: Tools for Building OLS Regression Models* (R Foundation for Statistical Computing, 2024).
80. Hamner, B. & Frasco, M. *Metrics: Evaluation Metrics for Machine Learning* (R Foundation for Statistical Computing, 2024).
81. Burgin, J. et al. The European nucleotide archive in 2022. *Nucleic Acids Res.* **51**, D121–D125 (2023).
- français Paul-Émile Victor' (IPEV) and the 'Muséum national d'Histoire Naturelle' (MNHN) in the context of the research project 'REVOLTA 1124', led by C.G. and M.E. The Regional Council of Brittany, the General Council of Finistère, the urban community of Concarneau Cornouaille Agglomération and the European Regional Development Fund (ERDF) are acknowledged for the funding of the MiSeq of the Concarneau Marine Station. We thank the Service de Systematique Moleculaire (UAR 2700 2AD MNHN, CNRS) and its team in Paris and Concarneau for access to the lab and reagents. The ARMS-MBON network is funded by the infrastructure programmes ASSEMBLE Plus (grant no. 730984) and the European Marine Biological Resource Centre, EMBRC. We thank the following people and observatories for the sample collection: Roman Wenne (Svalbard), Thierry Comtet/Stephane Loisel/Laurent Leveque/Frederique Viard (Roscoff), Liraz Levy (Eilat), Peter A. Staehr (Laesoe, Limfjord), Vasilis Gerovasileiou/Thanos Dailianis/Eva Chatzinikolaou/Giorgos Chatzigeorgiou/Kleoniki Keklikoglou/Emmanouella Vogiatzi (Crete). Funding for ARMS observatories in Norway was provided by the INTERREG project GEANS (North Sea Program of the European Regional Development Fund of the European Union). D.F. and S.S. are supported by the National Biodiversity Future Center (NBFC) of the National Recovery and Resilience Plan (NRRP) of Italy, Award Number: CN_00000033. This study has been conducted using E.U. Copernicus Marine Service Information;10.48670/moi-00021. This paper is a contribution to the SCAR-ANTOS Expert Group (<https://scar.org/science/cross/antos>) and an Italian contribution to the CCAMLR CONSERVATION MEASURE 91-05 (2016) for the Ross Sea region Marine Protected Area, specifically, addressing Annex 91-05/C ('long-term monitoring of benthic ecosystem functions'). We are very grateful to John K. Pearman for providing metadata and assistance with the processing of samples at the bioinformatic pipeline.

Author contributions

M.Cecchetto and S.S. designed and wrote the manuscript. S.S., M.Cecchetto, P.K., P.B., M.Chelchowski, M.M., A.P.K., M.Z., H.R., M.O., C.P., A.D., C.G., and M.E. performed the deployment and/or retrieval of ARMS structures and/or contributed to the design of the sampling procedures. M.Cecchetto performed the satellite data download and processing, all bioinformatic and statistical analyses. M.Cecchetto, S.S., G.F.F., and M.O. contributed to the design and/or review of the analyses. M.O., C.P., and K.E. performed the data collection and maintenance. S.S., G.F.F., and D.F. supervised the study, and all authors contributed to the revision and correction of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-52673-z>.

Correspondence and requests for materials should be addressed to Matteo Cecchetto.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements

The deployment, recovery and analyses for the ARMS deployed in Terra Nova bay were funded by the Italian National Antarctic Program (PNRA) projects '16_00120—TNB-CODE: Terra Nova Bay barCODing and mEta-barcoding of Antarctic organisms from marine and limno-terrestrial environments (PI: S.S.)' and '18_00078—B2—RosS-MODE: Monitoring of Ross Sea biodiversity by environmental DNA, barcoding, and meta-barcoding (PI: G.F.F.)'. The deployment and recovery of the ARMS deployed near Dumont D'Urville was supported by the 'Institut polaire

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024

¹Department of Earth, Environmental and Life Science (DISTAV), University of Genoa, Genoa, Italy. ²Institut de Systématique, Évolution, Biodiversité (ISYEB), Muséum National d'Histoire Naturelle, CNRS, SU, EPHE, UA, Paris, France. ³Institut de Systématique, Évolution, Biodiversité (ISYEB), Sorbonne Université, MNHN, CNRS, EPHE, UA Station Marine de Concarneau, Concarneau, France. ⁴Department of Marine Sciences, University of Gothenburg, Gothenburg, Sweden. ⁵Institute of Oceanology, Polish Academy of Sciences, ul. Powstańców Warszawy 55, Sopot, Poland. ⁶Nord University, Faculty of Biosciences and Aquaculture, 8049 Bodø, Norway. ⁷Department of Environmental Science and Policy, Università degli Studi di Milano, Milano, Italy. ⁸European Marine Biological Resource Centre (EMBRC-ERIC), Paris, France. ⁹Flanders Marine Institute (VLIZ), InnovOcean Campus, Jacobsenstraat 1, 8400 Oostende, Belgium. ¹⁰National Research Council of Italy—Water Research Institute (CNR-IRSA), I-28922 Verbania, Italy. ¹¹National Biodiversity Future Center (NBFC), I-90133 Palermo, Italy. ¹²Italian National Antarctic Museum (MNA, Section of Genoa), University of Genoa, Genoa, Italy.

✉ e-mail: matteocecchetto@gmail.com