

RESEARCH ARTICLE

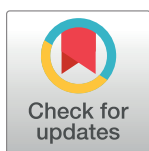
Biodiversity patterns of epipelagic copepods in the South Pacific Ocean: Strengths and limitations of current data bases

Manuela Pérez-Aragón^{1,2}, Ruben Escribano^{2,3}*, Reinaldo Rivera², Pamela Hidalgo^{2,3}

1 Doctoral Program of Oceanography, Facultad de Ciencias Naturales y Oceanográficas, Universidad de Concepción, Concepción, Chile, **2** Instituto Milenio de Oceanografía (IMO), Universidad de Concepción, Concepción, Chile, **3** Department of Oceanography, Facultad de Ciencias Naturales y Oceanográficas, Universidad de Concepción, Concepción, Chile

* These authors contributed equally to this work.

* rescribano@udec.cl



OPEN ACCESS

Citation: Pérez-Aragón M, Escribano R, Rivera R, Hidalgo P (2024) Biodiversity patterns of epipelagic copepods in the South Pacific Ocean: Strengths and limitations of current data bases. PLoS ONE 19(7): e0306440. <https://doi.org/10.1371/journal.pone.0306440>

Editor: Hans G. Dam, University of Connecticut, UNITED STATES OF AMERICA

Received: April 20, 2023

Accepted: June 18, 2024

Published: July 11, 2024

Copyright: © 2024 Pérez-Aragón et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All data used in this study on occurrences of Copepoda species and environmental variables are fully available in Zenodo repository <https://doi.org/10.5281/zenodo.8257222>.

Funding: This work was funded by the Instituto Milenio de Oceanografía (IMO) through the Agencia Nacional de Investigación Científica y Tecnológica de Chile (ANID) Grant IC_120019. The funders had no role in study design, data collection

Abstract

Basin-scale patterns of biodiversity for zooplankton in the ocean may provide valuable insights for understanding the impact of climate change and global warming on the marine ecosystem. However, studies on this topic remain scarce or unavailable in vast regions of the world ocean, particularly in large regions where the amount and quality of available data are limited. In this study, we used a 27-year (1993–2019) database on species occurrence of planktonic copepods in the South Pacific, along with associated oceanographic variables, to examine their spatial patterns of biodiversity in the upper 200 m of the ocean. The aim of this study was to identify ecological regions and the environmental predictors explaining such patterns. It was found that hot and cold spots of diversity, and distinctive species assemblages were linked to major ocean currents and large regions over the basin, with increasing species richness over the subtropical areas on the East and West sides of the South Pacific. While applying the spatial models, we showed that the best environmental predictors for diversity and species composition were temperature, salinity, chlorophyll-a concentration, oxygen concentration, and the residual autocorrelation. Nonetheless, the observed spatial patterns and derived environmental effects were found to be strongly influenced by sampling coverage over space and time, revealing a highly under-sampled basin. Our findings provide an assessment of copepods diversity patterns and their potential drivers for the South Pacific Ocean, but they also stress the need for strengthening the data bases of planktonic organisms, as they can act as suitable indicators of ecosystem response to climate change at basin scale.

Introduction

Spatial patterns of biodiversity in the marine environment have long been documented [1, 2]; however, factors and processes controlling such patterns remain unclear [3]. In the framework

and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

of conservation purposes and preserving ocean biodiversity, it becomes clear that the variation in the biogeographic regions from time series observations [4], as well as from geological time scale analyses [5], requires a better understanding of underlying mechanisms driving the spatial and temporal biodiversity patterns [6].

Zooplankton are a significant component of the marine ecosystem; they are the most widespread form of animal life on Earth, with the longest history of evolutionary continuity [7]. These organisms are also the foundation of life in the ocean, acting as the trophic link between primary producers and upper trophic levels [8], being important components of marine biogeochemical cycles [9]. They are also known to rapidly respond to oceanographic and environmental variations, such as fluctuations on temperature [10], oxygenation [11], acidification [12, 13], stratification [14, 15], primary production [16, 17], upwelling [10, 18, 19], circulation [17] and advection [20], thus providing themselves essential environmental sensors for a changing ocean.

Copepods are the taxon with the largest amount of accepted species described and documented in the World Register of Marine Species (WoRMS) [21]. It is also a well-represented taxonomic group in several data bases for marine species, such that a relatively thorough database on their distribution (horizontal and vertical) is available for the South Pacific Ocean [22]. Most copepod data bases only report species occurrences, or at higher taxonomic levels, and with emphasis on Calanoid copepods. Even that, such data can facilitate studies on the drivers controlling copepods species richness, how they are spatially structured, and the way they might respond to a changing environment. However, the study of large-scale copepod diversity patterns and their fluctuations driven by environmental forces have been scarcely conducted [23–25]. Moreover, open ocean studies on biogeography patterns of copepods are especially scarce in the Southern Hemisphere. For the vast South Pacific Ocean (SPO), that represents the largest yet the least-known marine ecosystem in the world, available data on copepods are mostly compiled in the Ocean Biodiversity Information System (OBIS) Portal [26]. An important task, however, to use these data sets for studies on copepods distribution and diversity, is the assessment of their quality in terms of resolution and coverage, both in time and space.

With respect to global distribution of plankton species richness, efforts have been made to test hypotheses explaining spatial patterns [23, 27]. For instance, the species-energy hypothesis [28, 29], stating that available energy can regulate species diversity, has been widely supported [30], but how diversity becomes linked to available energy is matter of debate [23]. In the same context, the best predictor of plankton diversity in the ocean appears to be temperature, exhibiting the greatest correlation or highest explanatory power as reported in many works [3, 23, 25, 30–35], although, when analyzing latitudinal patterns of copepod diversity in the Atlantic Ocean, Woodd-Walker et al. (2002) [36] found that species diversity was fundamentally related to variability in seasonality from the equator to polar regions rather than to temperature. The explanation was that more stable conditions and continuous primary production at low latitude can promote a higher diversity, while a highly variable and strongly seasonal cycle of primary production in high latitudes can favor the dominance of fewer species. In any case, the positive correlation between sea temperature and diversity of zooplankton continues to be a key issue for predicting large-scale spatial variability of plankton communities [37, 38]. Temperature is one of the two most important properties of seawater controlling its density and thus governing the dynamics and circulation of the water masses [39]. Temperature also plays a key role in the range of distribution of various marine organisms [34, 40]. For instance, copepods exhibit a strong dependence in their vital rates with temperature [17], since they are ectotherms, so that their population dynamics is strongly influenced by this factor, with major implications on their species composition and diversity patterns [18, 41].

Over a large-scale domain (e.g., a basin-scale), the role of temperature or other environmental correlates as biogeographic predictors for copepods is still unclear for the South Pacific region. The remaining questions are: do copepod species exhibit significant clustering or distinctive spatial assemblies? and, what are the most plausible environmental correlates to predict the diversity distribution of copepods at a basin scale? Lastly, is temperature alone the best predictor of such spatial structures, and whether such structures and biogeographic patterns replicate those observed in the North Pacific or Atlantic Ocean? In this study, we addressed these questions using a long-term (1983–2019) data base on occurrence of copepod species, extracted from the OBIS Portal, allowing us to further evaluate the robustness of these compiled data on species records of epipelagic copepods. For this, we assessed the diversity patterns of planktonic Copepoda and their relationship with oceanographic conditions in the upper 200 m of the western and eastern sides of the SPO to identify biogeographic regions and the presence of hot and cold spots of diversity. We aimed at testing the hypothesis that water temperature is the best predictor explaining spatial diversity patterns of planktonic copepods at basin scale in the SPO. Ultimately, our study will provide a first assessment of copepods diversity for the entire basin, helping us to gain insights into how such patterns may vary under ongoing ocean warming and, also allowing us to evaluate the robustness of current data bases for plankton organisms in this large ocean basin.

Materials and methods

Study area

The SPO possesses the greatest surface and seabed area, as well as the greatest volume, with a mean depth of 3993 m [42]. The ocean stretches from the Equator—including the islands of the Gilbert and Galápagos Groups, which lie to the northward thereof—until the parallel of 60°S, and between South America and the northeastern limit of the East Indian Archipelago (from New Guinea to the Equator), along the southern, eastern and northern limits of the Bismarck and Solomon Seas; the southeastern and northeastern limits of the Coral Sea; and the southern, eastern and northern limits of the Tasman Sea, going down the meridian of 146° 55'E starting at the South East Cape, the southern point of Tasmania, up to the parallel of 60°S [43, 44] (Fig 1). The general circulation in the SPO and major current systems are illustrated in Fig 1.

Data sources—Copepoda species and quality control procedures

Species occurrence records were obtained from the Ocean Biodiversity Information System, OBIS [22] (<http://www.iobis.org>) data base, using the *robis* [45] and *devtools* [46] packages implemented in the R Core Team (2021). The data were obtained for all occurrences between 0–200 m of the SPO, from 1993 to 2019. Following the retrieval of the data, we eliminated records without information on the geographic coordinates, coordinates equal to zero, or records located inside continents. We selected only records at level of species and excluded duplicate records. This procedure allowed us to compile data to estimate species diversity, but also resulted in a drastic reduction of information since many records are reported for genera or higher taxonomic levels (e.g. families), and in some cases without clear or wrong georeference. Taxonomy was revised and updated using the WoRMS portal (<http://www.marinespecies.org>) through the *match_taxa* function of *robis* package [45] implemented in R software (R Core Team, 2021). After we had cleaned and curated the data, a total of 22345 records of Copepoda were selected. The data on occurrences of Copepoda species are available in a Zenodo repository: <https://doi.org/10.5281/zenodo.8257222> (version 1). Finally, diversity indexes (see below) were generated based on biogeographic information of copepod biodiversity patterns,

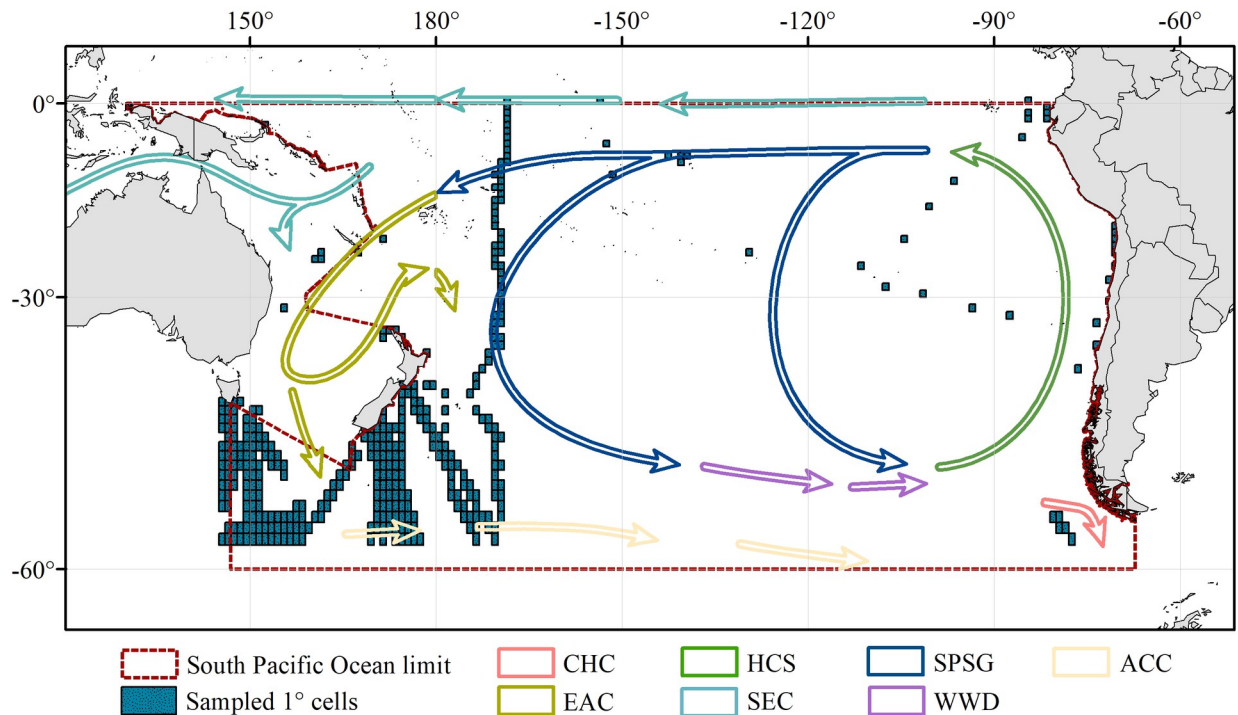


Fig 1. The South Pacific Ocean basin (red dashed area), its sampling coverage with georeferenced locations of the 393 cells (1° x 1°) which were generated (blue squares) from data spanning between 1993–2019. The major near-surface currents shown for the whole basin were obtained from Harvard Dataverse (<https://doi.org/10.7910/dvn/tkgo2z>). CHC = Cape Horn Current; EAC = East Australian Current; HCS = Humboldt Current System; SEC = South Equatorial Current; SPSG = South Pacific Subtropical Gyre; WWD = West Wind Drift; ACC = Antarctic Circumpolar Current. Map projection is WGS 84/PDC Mercator (EPSG 3832).

<https://doi.org/10.1371/journal.pone.0306440.g001>

implemented in QGIS 3.10 (QGIS, 2022) at 1 degree resolution, totaling 393 cells (Fig 1). However, for all the spatial analyses, the basin was divided into two sections where more data were available. i.e., eastern and western side, excluding a longitudinal band containing cells between -110° and -170°, that had low data coverage (only 16 cells available for that area) and could skew spatial patterns results. Therefore, 377 cells were considered for the analyses. Moreover, when representing alpha diversity patterns through kriging and an inverse distance weighted technique, all sampling cells containing less than 10 species were excluded, assuming they underestimated the species records and hence introduced biases in the patterns. Therefore, for those analyses, only 77 cells were considered.

Data sources—Environmental database and processing

The oceanographic variables used for analyses were obtained from Copernicus Marine Environment Monitoring Service (<https://marine.copernicus.eu>) to a resolution of 1 x 1 degree and 0.25 x 0.25 degrees. From this data source, we obtained sea surface temperature [47], salinity [47], total chlorophyll-a concentration [48], dissolved oxygen concentration [48], and mixed layer depth [49]. The resolution selected to analyze our data was 1 degree; hence, the environmental variables with a resolution of 0.25 degrees were resampled into a 1-degree resolution using QGIS 3.10 (QGIS, 2022). All data were averaged for the 0–200 m depth range, as well as from years 1993 to 2019. For temperature data, two additional variables were generated: temperature stability and standard deviation of temperature, which were calculated based on temperature data in temporal lags of two years, with the package *climateStability* [50] implemented

in R software (R Core Team, 2021). “Deviation” of temperature is defined as the mean standard deviation between time slices over time elapsed, which was calculated at the beginning and end of each of the two years temporal lags and dividing this result by the length of the interval to quantify deviation over time for each time slice, to finally average the result across all time slices. Later, “stability” was obtained when taking the inverse of the deviation and scaling it between 0 and 1 [50]. Therefore, a total of seven variables were included in our analyses and are available in Zenodo repository: <https://doi.org/10.5281/zenodo.8257222> (version 1). These environmental variables were preprocessed by standardizing them subtracting their mean and dividing them by their standard deviation for later performing variable selection by using Random Forests through the package VSURF [51] implemented in R software (R Core Team, 2021). This is a two-stage strategy based on a preliminary ranking of the explanatory variables using the random forests permutation-based score of importance [52]. We selected a subset of important variables from the second or the “interpretation” step according to the reduction of the out-of-bag error by adding them into the decision training models. These variables were later used for analysis of predictive models of species richness and species composition. Further, for the selected environmental variables, we calculated the Spearman rank-order correlation coefficient matrix to visualize their degree of association as well as check for redundancy.

Spatial analyses

To analyze the species composition and its spatial distribution, the most dominant species (most recurrent) were extracted from the data bases and then associated with major current systems of the SPO, assuming they were representing distinctive species assemblages from these currents. Also, for beta diversity and its components, turnover and nestedness, a presence-absence matrix (PAM) was generated through the package fuzzySim in R software (R Core Team, 2021) [53] from the list of species that occurred for each range of values. This PAM was later matrix plotted in PAST software [54], from which we selected visually the dominant species, i.e., the ones occurring in all ranges of values for each component.

Biodiversity hotspots were determined based on spatial-statistical criteria to describe biodiversity patterns for the 0–200 m depth water column. It was specifically aimed to detect cells or groups of cells (i.e., spatial clusters) with species richness greater than the richness expected in a random distribution. First, the evaluation of whether copepods depict spatial autocorrelation using Moran’s I statistic was conducted. Second, spatial hotspots were defined in ArcMap 10.4.1 (ESRI 2016) software using the High/Low Clustering (Getis-Ord G^*) statistic [55], which identifies spatial concentrations of an entity (in this case species richness per cell) or areas that contain higher/lower values than expected by chance for a given study area. For a hot spot, the observed G^* index indicated high values of richness clustered in the study area; whereas for a cold spot, the observed G^* index indicated that low values of richness were clustered in the study area.

Species richness (alpha diversity) is the absolute number of species living in a given area, giving equal weight to all resident species [56]. Species composition (namely beta diversity) can be viewed as a measure that compares inventory diversity at two different scales (alpha and gamma diversity). Beta diversity is also a measure of similarity between sites [57], and thus reflects the differences between local biological communities within a region. These differences arise from two phenomena which should be distinguished when understanding how biological diversity is distributed: nestedness and spatial turnover [58]. Nestedness of species occurs when the biota of sites with smaller numbers of species are subsets of the biota at richer sites. Contrary to nestedness, spatial turnover implies the replacement of some species by others as a consequence of environmental sorting or historical and spatial constraints [59]. An application of knowing these components of beta diversity is the attempt to use differences between

communities to determine biogeographical regions, as a community that is a subset of a larger one (nestedness) has no exclusive species, and thus such difference should not be taken into account to delimit regions; whereas by using indexes not affected by nestedness it is possible to delimit regions with unique biological communities [58]. We therefore calculated alpha diversity as species richness and beta diversity as species composition while distinguishing between turnover and nestedness. Alpha diversity was calculated in Biodiverse 3.1 software [60], whereas beta diversity was estimated using the packages betapart [61], CommEcol [62] and letsR [63] in R software (R Core Team, 2021), and using the following equation:

$$\beta_{sor} = \beta_{sim} + \beta_{sne} \equiv \frac{b+c}{2a+b+c} = \frac{b}{b+a} + \left(\frac{c-b}{2a+b+c} \right) \left(\frac{a}{b+a} \right), \quad (1)$$

where β_{sor} is Sørensen dissimilarity, β_{sim} is Simpson dissimilarity (i.e., turnover component of Sørensen dissimilarity), β_{sne} is the nestedness component of Sørensen dissimilarity, a is the number of shared species between two cells, b the number of species unique to the poorest site, and c the number of species unique to the richest site.

As biodiversity index, we used the Shannon-Wiener index which was calculated in Biodiverse 3.1 software [62], according to Eq 2 [64] in the following way:

$$H = - \sum_{i=1}^n p_i \cdot \ln p_i \quad (2)$$

where p_i is the number of samples (in this case, occurrences) of the i species as a proportion of the total number of occurrences in the neighborhoods ($1^\circ \times 1^\circ$ sampling cells). This proportion is estimated as:

$$p_i = \frac{n_i}{N} \quad (3)$$

where n_i is the number of records of the i species and N is the total number of records at species level in the sampled cell.

In order to account for the low sampling coverage and explore whether cells with low values of richness were causing biases in the analyses, we discarded the cells with values of alpha diversity < 10 , and later performed the spatial analysis using an inverse distance weighted (IDW) technique that interpolates a raster surface from points, limiting to the range of the values used to interpolate and giving more (or less) weight where there is more (or less) data. In addition, to explore further at higher taxonomic levels, we performed the same analysis considering genus and family levels. Both analyses were plotted through Kriging interpolation using ArcMap 10.4.1 software (ESRI 2016).

Since species richness in a sample is highly dependent on sample size or sampling efforts [65], and the ocean has under-sampled areas due to spatial biases in the distribution of sampling locations [6], a skew on the apparent patterns of marine species richness can be produced. Hence, redundancy statistics were used to count the total number of species, and an assessment of sample completeness was done for species richness through the R package iNEXT, which provides functions to compute and plot the seamless rarefaction and extrapolation sampling curves [65]. This latter analysis was performed for each main surface current system of the SPO, namely: Cape Horn current, East Australian Current, Humboldt Current System, South Equatorial Current, South Pacific Subtropical Gyre, and West Wind Drift, to evaluate whether some areas had a greater or lower coverage than others. The previously mentioned main surface current systems were downloaded from Harvard Dataverse [66].

We generated a biological-environmental spatial database, wherein each cell had the biogeographic information of copepod alpha and beta diversity together with oceanographic data

from the selected environmental variables. We employed Generalized Additive Models (GAMs) to evaluate the relationship between copepod alpha and beta diversity in the 0–200 m layer of the South Pacific Ocean and the environmental predictors while considering spatial autocorrelation in our analyses. Spatial correlation is a pattern in which observations are related to one another by the geographic distance between the observation, with locations close to each other exhibiting more similar values than those further apart, thus violating the assumption of independent and identically distributed errors, which can inflate type I errors [67]. This can lead to the selection of unimportant explanatory variables and poorly estimated parameters, moreover in ecology, as autocorrelation is a general property of ecological variables measured over geographic space [67]. Therefore, to take spatial autocorrelation into account, we used the residuals autocovariate (RAC) approximation [67] by including autocorrelation in the GAMs by adding another term to it that represents the influence of neighboring observations on the dependent variable at a particular location. Subsequently, the residuals of the auto-covariables were used as a new predictive variable in the models. The calculations of GAM+RAC were performed with the *spdep* package [68, 69] implemented in R software (R Core Team, 2021).

The model selection was done through Bayesian Information Criterion (BIC), that is based on information theory but within a Bayesian context and imposes a greater penalty for the number of parameters compared to the Akaike information criterion [70]. The best model is the one that provides the minimum BIC, denoted by BIC^* , and delta BIC can be computed as: $\Delta BIC = BIC_m - BIC^*$. This implies that for given M models, the magnitude of the delta BIC can be interpreted as evidence against a candidate model being the best model [70]. As a rule, values of ΔBIC less than 2 are not worth more than a mention, values between 2 and 6 indicate that evidence against the candidate model is positive, values between 6 and 10 indicate that evidence against the candidate model is strong, and values greater than 10 indicate that evidence against the candidate model is very strong [70]. The analyses were conducted through the package *MuMIn* [71], in R software (R Core Team, 2021).

Beta diversity was evaluated at each side of the SPO using Generalized Dissimilarity Modelling (GDM)-based spatial analysis technique, which involves mapping the first three axes derived from a principal component analysis (PCA) performed on the set of transformed predictor layers [72]. It first formulates the relationship between biological dissimilarity and spatial distance by transforming each predictor variable through 1-spline basis functions considering a term called ‘predicted ecological distance’ in a negative exponential function, with the assumption that dissimilarity between site-pairs increases monotonically from 0 to 1 and in a saturating manner with predicted ecological distance. Based on the maximum height of the spline function (and hence, the maximum value of the transformed predictor), which shows the importance of the predictor in explaining dissimilarities, only the predictors with maximum height over zero were selected for the generation and plotting of the spatial layers. A RGB (RedGreenBlue) color palette dimension was assigned to each of the three axes of the PCA, and they were mapped by combining them, finally enabling the visualization of areas predicted to have more similar species composition (shown as similar colors) or less similar species composition (shown as different colors). The GDM analysis was performed using the package *gdm* [73] in R software (R Core Team, 2021) and its plotting was done through Kriging interpolation using ArcMap 10.4.1 (ESRI 2016) software.

Results

Diversity patterns

For the South Pacific Ocean, a total of 178 species were reported at the upper 200 m water column of its basin. Regarding the main species present at each main surface current system, we

searched the dominant ones in terms of their occurrence along the whole period for each of the major current systems defined above. Most of the dominant species in the SPO were representatives of the Order Calanoida (S1 Table), except for the cyclopoid *Oithona similis*, which was the dominant species in the CHC, along with *Calanus simillimus* and *Pleuromamma robusta*. In the EAC, *O. similis* and *C. simillimus* were the dominant species accompanied by *Neocalanus tonsus*. In the SPSG *Calocalanus kristalli*, *Paracalanus parvus*, and *Calocalanus plumulosus* dominated; in the WWD *Calocalanus kristalli*, *Paracalanus parvus* and *Calocalanus plumulosus* were the most recurrent species; in the HCS *Paracalanus indicus*, *Oithona similis* and *Acartia tonsa* prevailed; and finally in the SEC *Paracalanus parvus*, *Calocalanus kristalli* and *Calocalanus plumulosus* were the most dominant ones (S1 Table).

For beta diversity and its components turnover and nestedness, the list of dominant species is shown in S2 Table. The dominant species for the mean beta diversity were the Calanoids *Acartia longiremis*, *Calocalanus kristalli*, *Calocalanus pavo*, and *Paracalanus parvus*, whereas turnover had similar dominant species, except by the substitution of *C. pavo* by *C. plumulosus*. Finally, nestedness resulted in a different species assemblage represented by *Calocalanus pavo*, *Euterpina acutifrons*, *Lucicutia flavicornis*, and *Mecynocera clausi*.

The spatial distribution of alpha diversity considering the complete data set with the curated data (i.e., 377 cells, excluding the -110°–-170° longitudinal band) is shown in Fig 2A. The removal of cells with values of alpha diversity < 10, left 77 cells located at the eastern and western sides of the SPO, that were used for the inverse distance weighted (IDW) technique analysis (Fig 2B), and for the Kriging interpolation (Fig 2C). Those results showed a similar tendency as the ones obtained with Kriging interpolation done to explore the outcome of this approach by using the complete dataset (393 cells, without excluding the -110°–-170° longitudinal band), i.e., showing lower values of diversity at the western side of the SPO with a peak in the tropical and temperate zone, decreasing towards higher latitudes, whereas the eastern side showed higher values peaking at the temperate area that decreased towards the Equator and from the coastal towards the oceanic zone (Fig 2B and 2C). When considering family level, the number of cells increased to 399, and the Kriging interpolation presented a pattern with more families observed towards the eastern tropical and equatorial zone, with lower values towards the temperate, subpolar, and western area; whereas its western side showed higher values of families towards the continent with peaks in equatorial and temperate areas, and decreasing towards the open ocean (Fig 2D). When considering genus level, the number of cells increased to 394, and the Kriging interpolation presented a similar pattern to that observed at species level for eastern alpha diversity in Fig 2A and 2B, with clearer peaks in northern Chile and off Ecuador; whereas its western side showed higher values of genus towards the continent with peaks in equatorial and temperate areas, and decreasing towards the open ocean (S1 Fig).

The Shannon-Wiener index of diversity for pelagic copepods exhibited a pattern that varied from that of alpha diversity. For the eastern and western sides of the basin, the higher values were within the moderate range (2.4 and 2.9, respectively), showing different spatial patterns. At the eastern side of the SPO, greater diversity is observed on a coastal band from the temperate zone towards the equator, with decreasing values towards the open ocean (Fig 3). At the western side, higher diversity was observed from subtropical areas towards temperate and subpolar zones, while also showing decreasing values towards the open ocean (Fig 3).

To identify the extreme areas in terms of diversity, we used the High/Low Clustering (Getis-Ord G) statistical analysis. This method yielded patterns consistent with those of diversity (in terms of species richness) and allowed us to identify most of the hot spots (high values of richness) present in western temperate-subpolar areas off Australia, and in the equatorial area off Peru and Ecuador, decreasing towards the south and into the open ocean (Fig 4).

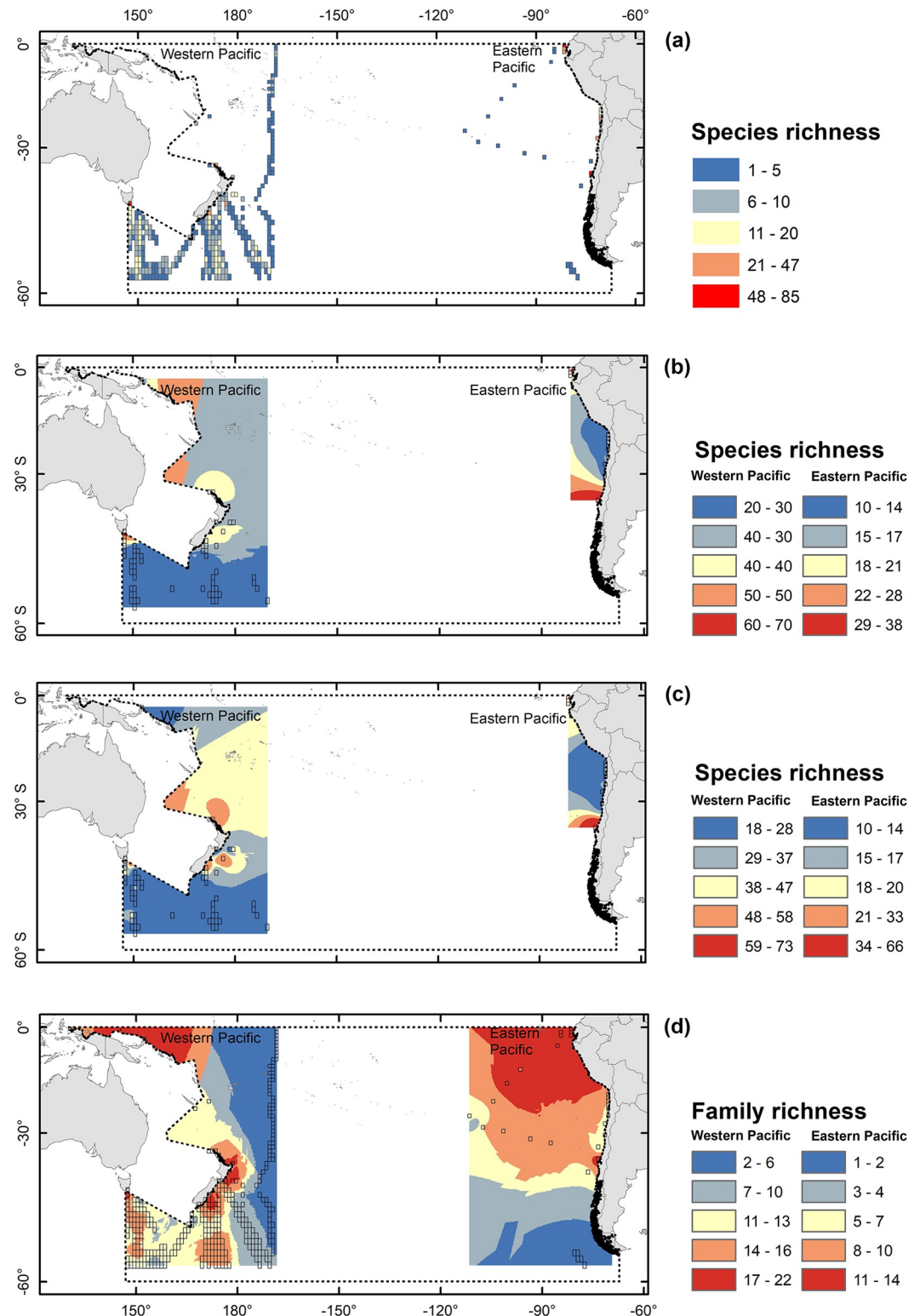


Fig 2. Richness patterns of Copepoda in the 0–200 m layer of the western and eastern side of the SPO for the period 1993–2019 based on species occurrence data obtained from OBIS portal. (a) Spatial distribution of alpha diversity considering the complete data cell with the curated data (i.e., 377 cells). **(b)** Inverse distance weighted (IDW) analysis for alpha diversity using cells with >10 species (i.e., only 77 cells from the total of 377 were selected). **(c)** Kriging interpolation for alpha diversity using cells with >10 species (i.e., only 77 cells from the total of 377 were selected). **(d)** Kriging

interpolation for family richness. Transparent squares are the 1° sampled cells used for interpolation, whereas the grey dotted line delimits the South Pacific Ocean. Map projection is WGS 84/PDC Mercator (EPSG 3832).

<https://doi.org/10.1371/journal.pone.0306440.g002>

The expected number of species, as a function of sampling units for different defined regions can illustrate the actual level of sampling coverage (Fig 5). As shown by species rarefaction curves, the sampling effort (samples per sampling units) varied extensively within and between regions, and no saturated curves were found in any of the regions, indicating a rather low coverage to assess the whole copepod community (Fig 5A). The distribution of the defined regions or current systems are illustrated in Fig 5B.

The spatial variability in diversity among potential distinct assemblies can be reflected in beta diversity, which may also indicate the degree to which such communities differ from each other. Beta diversity of copepods in the eastern SPO was higher in its equatorial-tropical area, with another area showing higher values at its temperate zone (Fig 6A); whereas the western side of the SPO had higher values of beta diversity expanding from tropical to subpolar areas (Fig 6A). This beta diversity (Fig 6A; mean beta = 0.579) was better explained by mean turnover (Fig 6B; mean turnover = 0.295) rather than mean nestedness (Fig 6C; mean nestedness = 0.284). The turnover rate at the eastern side of the SPO was higher in its equatorial-tropical area, with another area showing higher values at its temperate zone (Fig 6B); whereas the western side of the SPO had higher values of turnover expanding from tropical to subpolar areas (Fig 6B). Nestedness was found to be higher at the open ocean in the eastern side, with a higher values patch associated to the coastal area in the temperate zone (Fig 6C); whereas at the western side, it showed higher values in more oceanic areas at the temperate and subpolar zones (Fig 6C).

When applying the GDM-based spatial analysis to assess species composition, we found that salinity was the strongest predictor of the observed dissimilarities, followed by oxygen concentration, mean temperature and chlorophyll-a concentration. In contrast, mixed layer depth, temperature stability and standard deviation of temperature were the least important predictors, (S2 Fig). The spatial layers were generated and plotted with the best predictors and showed three clear latitudinal bands of similar color that kept almost the same distribution

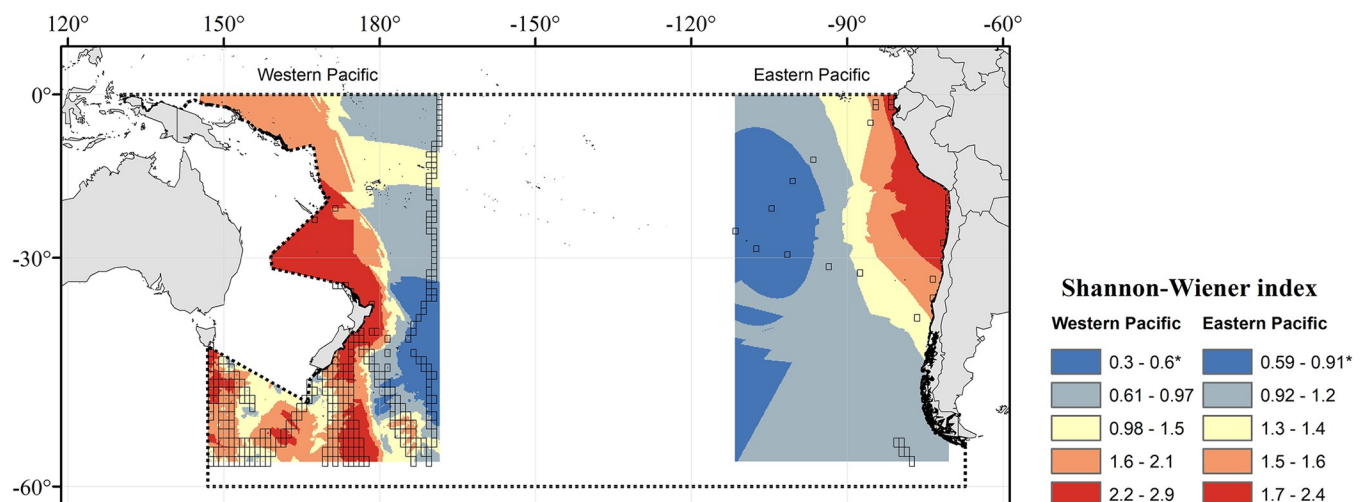


Fig 3. Shannon-Wiener index of Copepoda in the western and eastern sides of the South Pacific Ocean. Transparent squares are the 1° sampled cells used for Kriging interpolation, whereas the grey dotted line delimits the South Pacific Ocean. *Estimates based on available data at species level. Map projection is WGS 84/PDC Mercator (EPSG 3832).

<https://doi.org/10.1371/journal.pone.0306440.g003>

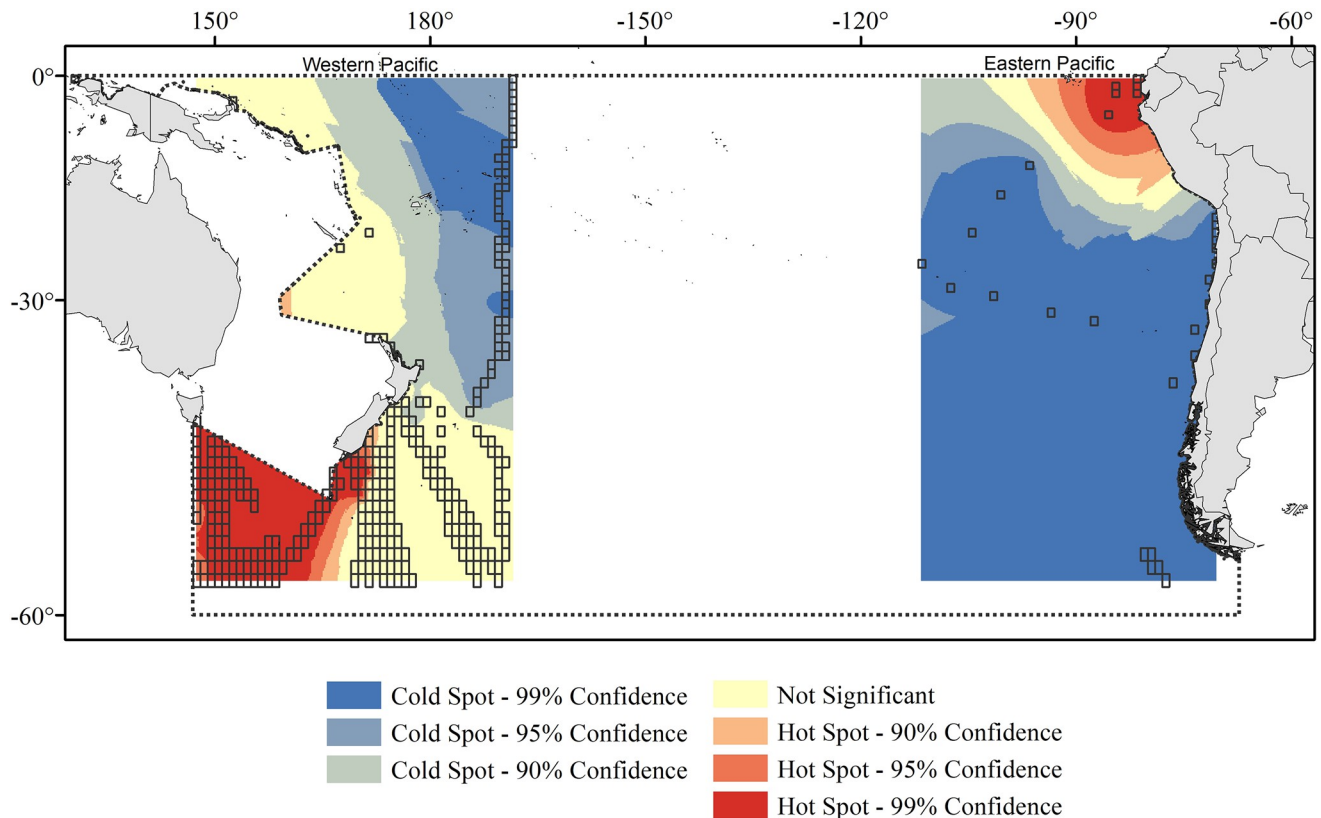


Fig 4. Getis-Ord G statistic of alpha diversity of Copepoda at the western and eastern sides of the SPO based on species occurrence data obtained from OBIS portal for the period 1993–2019. High and low values of richness are displayed in red and blue color, respectively. Color intensity denotes clusters' significance. Transparent squares are the 1° sampled cells used for Kriging interpolation, whereas the grey dotted line delimits the South Pacific Ocean. Map projection is WGS 84/PDC Mercator (EPSG 3832).

<https://doi.org/10.1371/journal.pone.0306440.g004>

when extracting the -110° – -170° longitudinal band, indicating a more similar expected composition of Copepoda or the existence of distinctive species assemblies (communities) with a variable degree of mix; however, they covered large areas over each side of the basin, indicating the existence of different communities (Fig 7). These distinct communities can also be defined in terms of their dominant species which are shown in S2 Table. Some of these dominant species are shared between current systems, although their presence also reflects a wide distribution over the SPO basin.

Data coverage, biogeographic patterns, and dominant species

This assessment was carried out in association with the major current systems illustrated in Fig 1. The general circulation of the SPO is dominated by the subtropical gyre, that has an elevated mean dynamic topography at its center [74] (near 15° to 20° S [75]). The South Equatorial Current represents the westward flow of the South Pacific Subtropical Gyre, that terminates entering the Coral Sea and bifurcating at the east of the Great Barrier Reef between 15° and 22° S [76] to form the southerly East Australian Current [77], that strengthens as it flows along the coast of Australia [77]. Southward of $\sim 33^{\circ}$ S, it begins to separate into filaments, forming the East Australian Current extension and the Tasman Front [77] that consists of a series of current jets flowing eastward, mainly between 33° and 35° S [78]. These appears to be a large component of subtropical water feeding into the South Pacific Subtropical Gyre

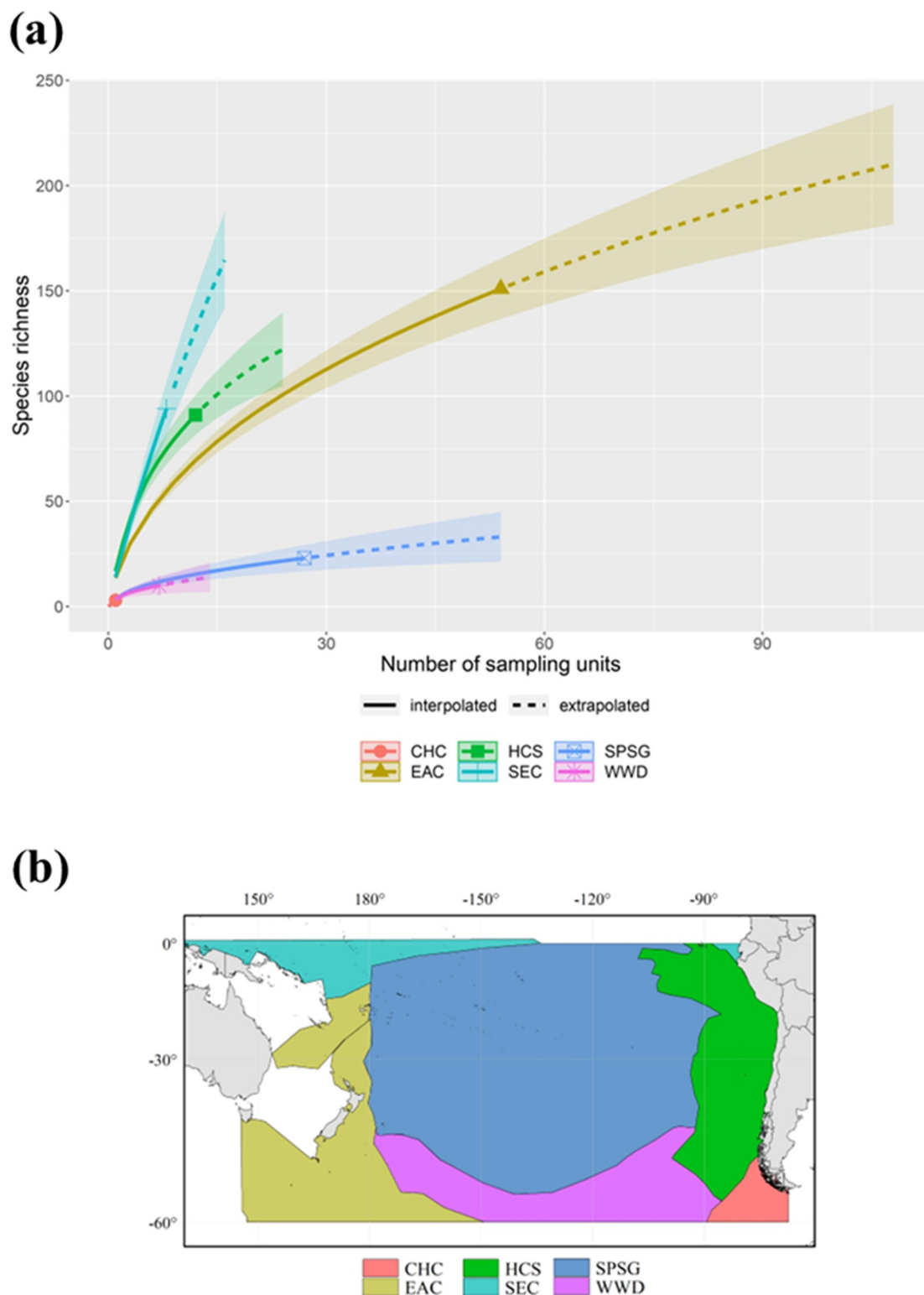


Fig 5. (a) Sample-size-based rarefaction (solid line segments) and extrapolation (dotted line segments) sampling curves for species richness with 95% confidence intervals (shaded areas) for Copepoda in the main surface current systems of the SPO. The symbols of each curve represent the reference samples. CHC = Cape Horn Current; EAC = East Australian Current; HCS = Humboldt Current System; SEC = South Equatorial Current; SPSG = South Pacific Subtropical Gyre; WWD = West Wind Drift. **(b)** Map of each main surface current system of the SPO.

<https://doi.org/10.1371/journal.pone.0306440.g005>

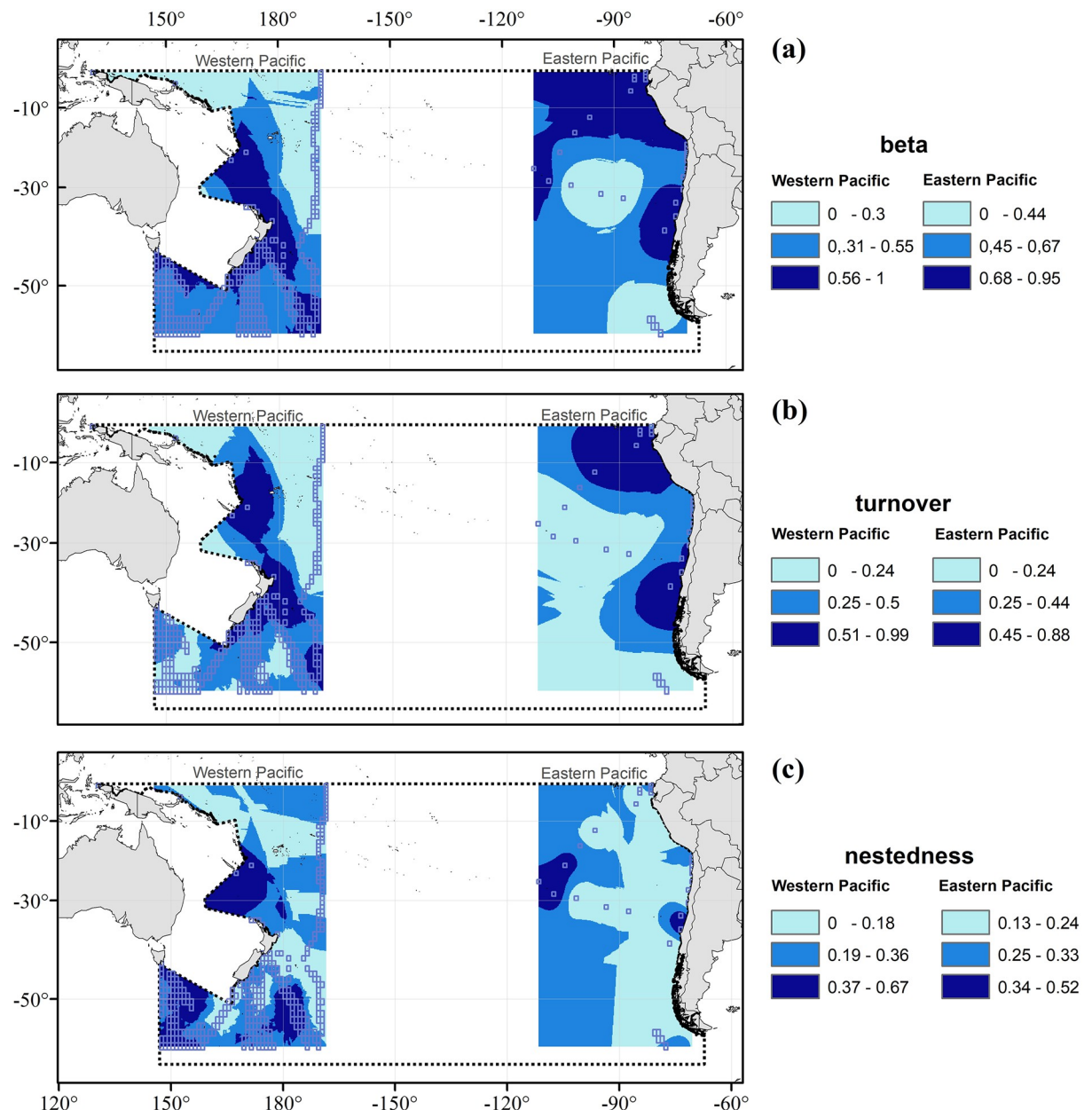


Fig 6. (a) Mean species composition and its components (b) turnover and (c) nestedness for Copepoda in the 0–200 m layer of the western and eastern sides of the SPO. Transparent squares are the 1° sampled cells used for Kriging interpolation, whereas the grey dotted line delimits the South Pacific Ocean. Map projection is WGS 84/PDC Mercator (EPSG 3832).

<https://doi.org/10.1371/journal.pone.0306440.g006>

circulation along the Sub-Tropical Front associated with the West Wind Drift, that streams equatorward of the Subantarctic Front associated with the Antarctic Circumpolar Current (ACC) [79, 80] and intersects the South American continent at ~45°S, where the Cape Horn Current and Humboldt Current System begin towards the south and the north, respectively [81, 82]. The latter is the largest of the four main Eastern Boundary Upwelling Systems—as the other three are embedded in the California, Canary, and Benguela Currents [83]—and is shaped by a broad current of fresh, cooler Sub-Antarctic Surface Water along the eastern rim of the subtropical gyre [81], extending from southern Chile (~45°S) to northern Peru and

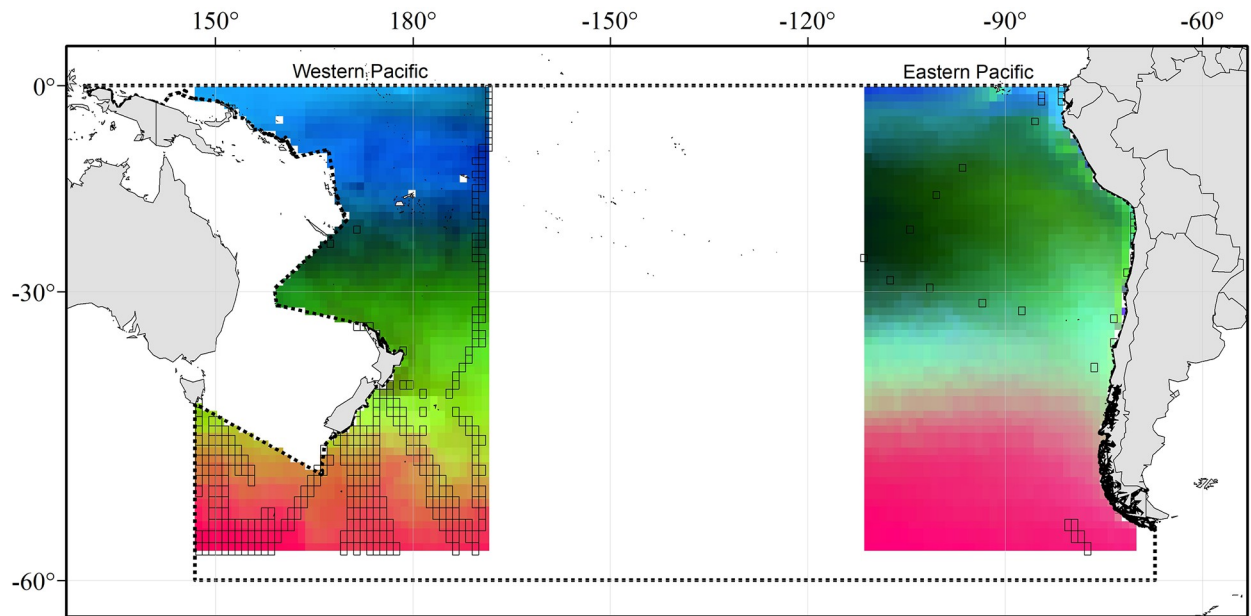


Fig 7. Predicted spatial variation in species composition of Copepoda at the western and eastern sides of the SPO based on a principal component analysis (PCA) of the Generalized Dissimilarity Modelling (GDM)-transformed environmental predictors. Colors represent gradients in species composition. Locations with similar colors indicate more similar expected composition. Dashed line at both sides of the antemeridian (180° longitude) delimits the South Pacific Ocean. Map projection is WGS 84/PDC Mercator (EPSG 3832).

<https://doi.org/10.1371/journal.pone.0306440.g007>

Ecuador ($\sim 4^{\circ}\text{S}$), where cool upwelled waters collide with warm tropical waters forming the Equatorial Front [81]. The South Pacific Subtropical Anticyclone or the South Pacific High, which spins counter-clockwise, is the most influential off the west coast of South America [74, 84] and acts as the dominant forcing of the subtropical gyre [85]. However, it presents seasonal variation, abiding at its northern position (26°S , 86°W) during the late austral fall and winter, when it is also closer to the South American continent and its intensity is weaker; whereas, during austral spring and summer the South Pacific High moves southwest (37°S , 108°W) and depicts its maximum intensity [85]. The latter (former) conditions generate stronger (weaker) Equatorward winds that favor (disfavor) coastal upwelling offshore central-south Chile [81]. In terms of sampling coverage for epipelagic copepods, only 4.02% of the total area of the epipelagic SPO has been sampled between 1993 and 2019. In fact, all the main current systems have been sampled over less than 2.5% of their area, each covering less than 1% of the total area of the SPO. An exception is the East Australian Current System with 17.28% of its area sampled and covering 2.76% of the total area of the SPO (Table 1).

Environmental data selection and predictive models. The routine VSURF selected five oceanographic factors to be tested for influence on species richness (arranged in decreasing order of importance): mean temperature (T_{mean}), salinity (S), dissolved oxygen concentration (O_2), chlorophyll-a concentration (Chla), and the mixed layer depth (MLD). For species composition, all seven oceanographic variables were selected (arranged in decreasing order of importance): mean temperature (T_{mean}), chlorophyll-a concentration (Chla), dissolved oxygen concentration (O_2), standard deviation of temperature (T_{sd}), salinity (S), temperature stability (T_{stab}), and mixed layer depth (MLD). The Spearman correlation matrices (S3A and S3B Fig) obtained for each subset of variables depicted that all were weakly correlated according to a Spearman's rank correlation coefficient [$\rho < 0.7$], with exception of T_{mean} and O_2 , which showed $\rho = -0.99$.

However, in addition to environmental variables, data on Copepoda richness within the upper 200 m of the SPO depicted that the species and assemblies were significantly associated

Table 1. Summary of spatial characteristics and cells sampled per each main surface current system of the SPO. CHC = Cape Horn Current; EAC = East Australian Current; HCS = Humboldt Current System; SEC = South Equatorial Current; SPSG = South Pacific Subtropical Gyre; WWD = West Wind Drift; n.a. = not applicable.

Current	Area (Km ²)	Total number of cells	Number of sampled cells	Total number of observations	Number of species observed per current	Sampled area per current (Km ²)	% of the current's area sampled	% of the area of the SPO sampled
CHC	2.65 *10 ⁶	152	5	55	3	6.16 *10 ⁴	2.32	0.05
EAC	1.93 *10 ⁷	1468	270	18207	148	3.33 *10 ⁶	17.28	2.76
HCS	1.59 *10 ⁷	1230	16	2527	91	1.97 *10 ⁵	1.24	0.16
SEC	6.74 *10 ⁶	685	12	510	94	1.48 *10 ⁵	2.19	0.12
SPSG	6.14 *10 ⁷	4436	67	899	24	8.26 *10 ⁵	1.34	0.68
WWD	1.46 *10 ⁷	804	23	147	10	2.83 *10 ⁵	1.94	0.24
Total	1.21 *10 ⁸	8775	393	22345	n.a.	4.84 *10 ⁶	n.a.	4.02

<https://doi.org/10.1371/journal.pone.0306440.t001>

over the space according to Moran's I statistics. This is shown by regression residuals which exhibited significant spatial autocorrelation, i.e., there are spatial-dependent elements in the regression residuals ($I = 0.012$, p -value = 0.015). Therefore, the residuals autocovariate (RAC) was added as a new variable for the model selection, after which the spatial autocorrelation according to Moran's I statistics was not significant ($I = 0.003$, $p = 0.091$).

When looking for environmental predictors to explain the distribution of species richness, the best GAM explaining species richness included Chla, S, T_{mean} and RAC (Table 2). The adjusted R^2 value revealed that 60% of the variability is explained by this model. The predicted values of alfa diversity were concordant with the observed data ($r = 0.77$; S4 Fig).

Regarding species composition within the 0–200 m range, a significant spatial correlation according to Moran's I statistics was observed, as regression residuals exhibited significant spatial autocorrelation, ($I = 0.018$, p -value = 0.005). After we added RAC as a new variable for the model selection, spatial autocorrelation did not occur based on Moran's I statistics ($I = -0.006$, p -value = 0.726).

The best GAM model explaining the species composition of copepods in the SPO included Chla, O₂, S, T_{mean} and RAC (Table 3). The adjusted R^2 value revealed that 70% of the variability is explained by this model. The predicted values of beta diversity were concordant with the observed data ($r = 0.84$; S5 Fig).

Discussion

Spatial patterns of biodiversity and their predictors

Hot and cold spots patterns of copepod distribution in the SPO do not seem to follow the previously reported latitudinal trends of global patterns of marine species with peaks of diversity

Table 2. GAM+RAC models for species richness. Statistics acronyms are BIC = Bayesian information criterion, Δ BIC = delta BIC (i.e., the difference in BIC score between the best model and the model being compared). Predictors' acronyms are: T_{mean} = mean temperature, S = salinity, Chla = chlorophyll-a concentration, MLD = mixed layer depth, autocovariate = residuals autocovariate. The best model is highlighted in bold.

	Model	BIC	Δ BIC
1	Species richness ~ Chla, S, T_{mean} + autocovariate	2364.54	0
2	Species richness ~ Chla, MLD, S, T_{mean} + autocovariate	2365.47	0.93

<https://doi.org/10.1371/journal.pone.0306440.t002>

Table 3. GAM+RAC models for species composition. Statistics acronyms are BIC = Bayesian information criterion, Δ BIC = delta BIC (i.e., the difference in BIC score between the best model and the model being compared). Predictors' acronyms are: T_{mean} = mean temperature, T_{stab} = temperature stability, S = salinity, Chla = chlorophyll-a concentration, DO_2 = dissolved oxygen concentration, MLD = mixed layer depth, autocovariate = residuals autocovariate. The best model is highlighted in bold.

	Model	BIC	Δ BIC
1	Species composition ~ Chla, DO_2, S, T_{mean} + autocovariate	553.32	0
2	Species composition ~ Chla, MLD, DO_2 , S, T_{mean} + autocovariate	551.84	1.48
3	Species composition ~ Chla, MLD, DO_2 , S, T_{mean} , T_{stab} + autocovariate	551.68	1.64

<https://doi.org/10.1371/journal.pone.0306440.t003>

at subtropical latitudes and a gradual decrease towards temperate and polar regions [23, 34, 86–90]. However, the presence of hot spots equatorial area off Peru and Ecuador (Fig 4) coincides with the Eastern margin biogeographic region where coastal (0–250 Km offshore) and eastern (250–1000 Km offshore) boundary currents that originate from the Humboldt Current and has high levels of production [91]; whereas the hotspots observed in western temperate-subpolar areas off Australia (Fig 4), corresponds, according to their zoning, to the South Subtropical Convergence and to the Chatham Rise zones, both frontal zones that marks the intersection of colder sub Antarctic waters with warmer tropical waters thus supporting substantial production [92]. The latitudinal trend, suggesting a plateau distribution of diversity from tropical to temperate regions, can be observed at the spatial distribution of the Shannon index shown in Fig 3, which shows greater diversity over a coastal band from the Equator towards the temperate area in the eastern zone, and higher diversity from subtropical areas towards temperate and subpolar zones at the western zone. Shannon Wiener, however exhibited some discrepancies with species richness. For instance, when comparing the northern vs. the southern portion of the eastern SPO (i.e. the Humboldt Current). In this regard, it is important to consider that we estimated the Shannon Wiener based on the species occurrences over the study period, not on species abundance, whereas species richness represents the total number of species throughout the study period, independently of their frequency of occurrence. Therefore, Shannon Wiener and species richness may show different patterns of diversity, when comparing regions where copepod populations are highly frequent, since their populations are more continuously present during their annual cycles, such as in tropical or subtropical regions, or they have very seasonal life cycles, with low frequency of occurrence, such as in temperate regions [93].

Regarding latitudinal trends, Woodd-Walker et al. (2002) [36] suggested that the decreasing trend with latitude may arise from a more stable seasonal cycle of productivity in the tropical/subtropical area supporting a stable high-diversity community in comparison with a strongly seasonal productivity in temperate/polar regions which may limit the number of copepods species and so resulting in a low diversity pattern. This latitudinal pattern is reflected in our spatial distribution of Shannon-Wiener index in both sides of the SPO, but it is not clear for the spatial distribution of alpha diversity, that at its western side showed lower values of diversity and had a similar tendency with a peak in the tropical and temperate zone, decreasing towards higher latitudes, although the eastern side showed higher values peaking at the temperate area that decreased towards the Equator and from the coastal towards the oceanic zone.

The zonal pattern on the other hand, remains with much uncertainty due to a poor sampling coverage in the oceanic region of the SPO. Our study shows that maximum values of copepod richness occur around tropical-subtropical areas in both eastern and western sides, but with a strongly reduced diversity in the oceanic region in areas with too few sampled cells. In this regard, Rombouts et al. (2009) [23] suggested that temperature-predicted diversity of copepods should be higher in a subtropical band in both hemispheres, although with no

contrasting differences between coastal and oceanic regions. Such patterns have also been found in other planktonic groups, such as euphausiids [34], tintinnid ciliates [94], foraminifera [34, 95]; as well as benthic marine invertebrates, such as prosobranch gastropods [32]; and higher trophic level organisms, fishes [34, 96], sharks [34], squids [34] and cetaceans [34]. For instance, in a recent work [19] a total number of 121 species of copepods was reported from a single cruise carried out in the South Pacific Subtropical Gyre, indicating this area can be comparable to the North Pacific Subtropical Gyre in which 125 species were reported by McGowan and Walker (1979) [97], although more recently Vereshchaka et al. (2017) [98] indicated that more than 240 species can be found in the central-south Pacific region. Therefore, the species records available at OBIS may not represent the potentially rich copepod community in the central gyre of the SPO; and that is why we opted for dividing the basin into a eastern and western side and exclude the longitudinal band between -110° – -170° for performing our analyses, as this area had too few sampling.

Temperature and its stability were the main factors, initially proposed in our hypothesis, as predictors of copepod biodiversity patterns in this basin. However, hot spots were also found off Chile, Peru and Ecuador along the main coastal upwelling system, which might be associated to 1) the habitat features hypothesis [34] related to ecosystem size and mesoscale processes, that are known to influence positively copepod diversity [34, 86, 90]; and 2) the productivity-richness hypothesis that predicts a positive effect of primary productivity on richness [34], supported by other studies finding a positive relationship between copepod diversity and chlorophyll-a concentration [86, 87, 89, 99]. Biogeographic patterns can also be explained on the basis of evolutionary [100, 101] and spatial processes [102] which are considered as the historical context for observed spatial distribution. Nevertheless, as stressed above, limited data coverage precludes a clear conclusion on underlying mechanisms explaining observed patterns of diversity.

Considered the above-mentioned potential mechanisms, we selected environmental variables to evaluate factors explaining copepod species richness and species composition. All variables were correlated to each other with a Spearman's rank correlation coefficient of $|\rho| < 0.7$, except for T_{mean} and O_2 , that showed $\rho = -0.99$. In general, variables with correlation values of $|\rho|$ less than 0.7 are not recommended for selection [103]; nevertheless, despite this high value of $|\rho|$, we chose to include both variables instead of conducting further testing. This is because dissolved gases in the ocean tend to decrease their concentration with higher temperature [104], and so this relationship would not reflect a systematic bias. The oceanographic factors explaining copepod species richness ($Chla$, S , T_{mean}) and species composition ($Chla$, O_2 , S , T_{mean}) in the SPO are of significance in a lesser or greater degree at different zones of the study area; and, they have also been documented to be of importance on a macroscale basis for copepod abundance and body size [25]. Moreover, the GDM-based spatial analysis conducted for evaluating species composition showed that the stronger predictors of the observed dissimilarities were the same variables (S , O_2 , T_{mean} , $Chla$; S3 Fig).

It has been found that chlorophyll strongly correlates with copepod abundance in the ocean [34, 86, 99, 105]. Oxygen is also a significant factor influencing copepod distribution in the ocean, because its availability is a major driver for temperature-size responses in aquatic organisms [106], and the variation in its concentration, for example, due to deoxygenation, can affect copepod physiology, with drastic impacts in both coastal [11, 107] and oceanic [108] species. Warmer water increases copepod growth, development and molting rates [107], with an associated increase in the oxygen demand. This can lead to a reduction in population growth by limiting the hatching success and naupliar growth in early life stages [11], as well as reducing egg production and somatic growth in adults [107], together with shifts in their depth distribution [108]. This involves physiological adaptations that may vary according to

the species, their body weight and extent of motion [109]. Altogether, these combined drivers have a potential effect on copepod ecology and, consequently, on their biogeographic patterns.

Salinity and its variation also affects copepod distribution [110], biomass [111], reproduction [112, 113], growth [114, 115], body size [25] or fatty acid synthesis [115], with various ranges of tolerance that may or may not be favorable, according to their biology and location. This may be relevant since global warming in some areas of the ocean can lead to an increase in salinity (evaporation or greater circulation) or a reduction of the same (greater run-off of fresh water into the ocean, rainfall, or ice melt), producing osmotic stress that may be costly for their physiology.

Temperature itself plays a fundamental role in the marine ecosystem, since it controls physiological rates [116], metabolism [116], body size [25], reproduction [10], mortality [117] and community structure [24, 118], even at large spatial scales [25, 35, 41, 119]. Thus, it may be thought that temperature also influences the distribution of organisms in the ocean, as it can substantially vary in its three dimensions. Over the meridional axis, temperature exhibits a strong gradient from the equator to the poles, with a similar gradient commonly present from coastal areas to central gyres, whereas in the vertical axis, it has a strong temperature gradient from the upper mixed layer down to the deep, cold waters. These gradients in the water column structure may contribute to the generation of large-scale spatial trends in the distribution of organisms such as copepods [120]. Although this study did not separate copepods into functional groups according certain traits more specifically linked to environmental variations, the fact that composition changed mainly due to oceanographic factors that influence the nesting of functional groups, highlights and support those findings [121], as beta diversity is a proxy of difference in number of species among ecosystems, thus reflecting changes or gradients in the environment as well.

Species distribution and the structure of species assemblages may additionally obey some physical processes (e.g., large-scale circulation, currents, eddies, fronts, interaction with the atmosphere) and the effect of local conditions. In this sense, it is possible that in addition to more direct effects on copepod populations, temperature is reflecting variability in water masses distribution controlled by large-scale circulation. The distribution of drifting plankton, including copepods, is largely affected by near-surface currents, allowing species radiation and colonization processes [122], and so influencing diversity patterns.

Regarding the structure of the copepod community, we found that some dominant species were representing different current systems, and thus distinct plankton communities. However, we could only assess most occurrences of species (not abundance) for each main surface current system of the SPO (S1 Table). The occurrences distribution of dominant species seemed to be concordant with patterns documented for copepod diversity on a large scale, with cyclopoids showing greater diversity in the subtropics and calanoids in the temperate zone [25, 123]. Dominant species in the each of the major currents are mostly represented by small Calanoid copepods and the widely distributed Cyclopoid *Oithona similis*. Some of the recurrent species are *Paracalanus parvus* and *P. indicus*. Both species are possibly a single species traditionally known as *P. parvus*, but morphologically close to *P. indicus*, although most likely being a new, non-described species [124]. Lately, it has been referred as *P. cf. indicus* in the Eastern SPO [125]. The other recurrent species are *Calocalanus* spp. They are also known as small-sized copepods and mostly found in oceanic waters [126]. The high occurrence of small-sized oceanic copepods, and low presence of large-sized Calanoids, such as *Calanus* spp., *Calanoides* spp., and *Pleurommama* spp., which are abundant in coastal areas of the SPO [38, 93] may indicate that small copepods prevail, or are numerically more abundant than large-sized ones. Also, most data sets do not include harpacticoids or other cyclopoid copepods, such as *Oncaea* spp. which may be highly abundant in oligotrophic waters of the SPO [127].

The OBIS data base does not include the observed number of species in each of the surface current systems (see [S1 Table](#)) when compared to total species inventories. For example, the maximum species richness available in OBIS for the East Australian Current (EAC) and the Humboldt Current System (HCS) is 148 and 91, respectively (see [Table 1](#)); that account for ca. a quarter of the species inventory compiled by the Banyuls Observatory data base on copepods for those areas [[128](#)]. In the other current systems, similar levels of differences between OBIS data and species inventories can be applied.

Current data limitations over space and time

Sampling biases must be considered when validating the robustness of the derived spatial patterns and the role of environmental drivers. For example, it is crucial to consider that due to spatial gaps of data ([Fig 1](#), [Table 1](#)), our results may be strongly constrained by sampling coverage [[129](#), [130](#)].

The spatial sampling effort was evaluated over the SPO through rarefaction/extrapolation curves by grouping the cells into the main surface current systems of the SPO ([Fig 5A](#)). It was found that species richness counts did not reach a clear asymptote at any of the defined current systems. Thus, the SPO is under-sampled for copepods. This aspect may have had an influence on Shannon-Wiener index values ([Fig 3](#)), as moderate to high values (from 1.7 on) were estimated only for a small area of the SPO. Subsequently, large areas exhibiting much lower values of richness and biodiversity may result from lack of sampling, description, and digitization of copepod data rather than having a relatively low diversity of species on a significant part of its extension. Moreover, the GDM-based spatial analysis done for evaluating species composition was performed with ecological data based on occurrence instead of abundance, which requires the assumption that any location where the species has not been observed can be treated as absence; therefore, making under-sampled locations more problematic [[72](#)]. Although invertebrates are the most common group documented after plants [[131](#)], and that an increasing number of data papers over time reflects progress in digitization and online platforms development for reporting observations that serve for biodiversity research [[131](#), [132](#)], it has been estimated that less than 10% of specimens' registers are digitized [[133](#)], with even less data available online [[134](#)]. The lack of description and cataloguing of species and the incomplete knowledge regarding their geographic distribution (i.e., Linnean and Wallacean shortfalls, respectively) are among the most common setbacks for biodiversity research using databases [[130](#)], representing a practical limit for biodiversity knowledge that may lead to misidentification of ecological processes [[130](#)]. We further addressed this issue by estimating redundancy index, that indicated a middling sampling over the SPO ([S6 Fig](#)) with a lower coverage on its eastern side, that highlights the gaps on great extensions and depicts higher values for areas where a greater sampling effort was seen. Moreover, through Moran's I index, it is possible to infer that the distribution of richness is more spatially clustered than expected if underlying spatial processes were random, as biodiversity aspects of species and communities are linked to complex interactions between physical, chemical, and biological factors. Considering this, and the developments in statistical modeling making possible to account for biases within heterogeneous data [[135](#)], we chose to use the RAC approximation [[67](#)] for the GAM analyses. This allows the incorporation of another variable into the models to take spatial autocorrelation into account and maintain strong predictive and inferential performance [[67](#)]. However, despite the usefulness of these methodologies, more effort is needed for improving the number and quality of distribution records used as primary data in macroecological studies [[136](#)], with databasing focused on georeferencing the information obtained from data collections [[129](#)].

Copepod resilience to climate forcing

Copepod diversity may have remained stable over a long time and large space due to a strong resilience to environmental variability. In this aspect and when dissecting species composition into turnover and nestedness, we observed that the former was higher in areas where hot spots were also identified (Figs 6 and 4, respectively), coinciding with the flow of cold currents (Humboldt current, Zeehan Current in southern Australia, and southern part of the East Australian current off New Zealand into the West Wind Drift). Therefore, basin-scale ocean currents seem to be an important oceanographic driver explaining the distribution of copepods in the SPO. The potential of large-scale ocean currents generating distributional patterns at various levels in the ocean has recently been studied more across different taxa [137–141], showing results that support the hypothesis and suggest that it may be a general process transcending taxa and spatial scales [141]. All these large-scale circulation patterns may allow the recolonization and mixing of copepod assemblies, thereby acting as a buffer to cope with the local impact of warming, deoxygenation, or other altered environmental conditions driven by natural perturbations or anthropogenic origin. This buffering effect may allow planktonic organisms such as copepods to exhibit a high resilience to large-scale changes in hydrographic conditions due to global warming.

Conclusion

The diversity distribution of copepods in the SPO follows general patterns observed in other organisms in the ocean, although patterns over a basin scale with presence of hot and cold spots do not replicate observed patterns reported for the North Pacific or Atlantic Ocean. The key environmental correlates for such biodiversity patterns were found to be mean temperature, chlorophyll-a concentration, salinity, oxygen concentration, and residuals autocovariate. A significant role for causing and maintaining copepod diversity patterns was also attributed to large-scale circulation processes, which may act as a buffer for changes in local conditions and allow species recolonization. Therefore, we suggest that planktonic copepods may exhibit a strong resilience to climate change impact.

However, caution should be taken in interpreting our findings since our modeling approaches and spatial analyses were strongly constrained by spatial and temporal gaps in sampling efforts and data availability. Future analyses are expected to improve observed patterns with the completion of global biodiversity data bases.

Supporting information

S1 Fig. Kriging interpolation analysis done for the western and eastern side of the SPO with 394 cells accounting for number of genera. Transparent squares are the 1° sampled cells used for Kriging interpolation, whereas the grey dotted line delimits the South Pacific Ocean. Map projection is WGS 84/PDC Mercator (EPSG 3832). (JPG)

S2 Fig. Maximum height of the spline function (hence the maximum value of the transformed predictors), indicating the stronger predictors of the observed dissimilarities for the Generalized Dissimilarity Modelling (GDM)-based spatial analysis. Their acronyms are: s = salinity, o2 = dissolved oxygen concentration, tmean = mean temperature, chl-a = chlorophyll-a concentration, mld = mixed layer depth, tstab = temperature stability, tsd = standard deviation of temperature. The spatial layers were generated and plotted with the predictors with heights over zero (i.e., s, o2, tmean and chl-a). (JPG)

S3 Fig. Spearman correlation matrices of environmental variables used in GAM models for (a) alpha diversity and (b) beta diversity. Positive correlations are displayed in red and negative correlations in blue color. Color intensity is proportional to the correlation coefficients. In the right side of the correlogram, the legend color shows the correlation coefficients and the corresponding colors. Their acronyms are: Tmean = mean temperature, Tstab = temperature stability, Tsd = standard deviation of temperature, S = salinity, Chla = chlorophyll-a concentration, O2 = dissolved oxygen concentration, MLD = mixed layer depth.
(JPG)

S4 Fig. Correlation between observed and predicted species richness. Pearson's r coefficient: 0.77 (p -value < 0.01). Negative residuals (below the reference line) indicate knowledge shortfalls, whereas positive residuals (above the reference line) indicate underestimated species richness.
(JPG)

S5 Fig. Correlation between observed and predicted species composition. Pearson's r coefficient: 0.84 (p -value < 0.01). Negative residuals (below the reference line) indicate knowledge shortfalls, whereas positive residuals (above the reference line) indicate underestimated species composition.
(JPG)

S6 Fig. Redundancy index obtained for the western and eastern side of the SPO. Values close to 1 indicate good sampling, whereas values close to 0 indicate poor sampling. Transparent squares are the 1° sampled cells used for Kriging interpolation, whereas the grey dotted line delimits the South Pacific Ocean. Map projection is WGS 84/PDC Mercator (EPSG 3832).
(JPG)

S1 Table. Dominant copepod species in terms of occurrence for the main surface current systems of the South Pacific Ocean. CHC = Cape Horn Current; EAC = East Australian Current; HCS = Humboldt Current System; SEC = South Equatorial Current; SPSG = South Pacific Subtropical Gyre; WWD = West Wind Drift.
(DOCX)

S2 Table. Dominant copepod species of the South Pacific Ocean in terms of occurrence for all ranges values of beta diversity and its components turnover and nestedness.
(DOCX)

Acknowledgments

We are grateful to all the scientific community providing data on marine biodiversity to OBIS portal, and special thanks to ESPOBIS, the Eastern South Pacific Node of OBIS. Two anonymous reviewers greatly helped to improve our work.

Author Contributions

Conceptualization: Manuela Pérez-Aragón, Ruben Escibano, Reinaldo Rivera.

Data curation: Reinaldo Rivera, Pamela Hidalgo.

Formal analysis: Manuela Pérez-Aragón.

Funding acquisition: Ruben Escibano.

Investigation: Manuela Pérez-Aragón, Ruben Escibano, Reinaldo Rivera, Pamela Hidalgo.

Methodology: Manuela Pérez-Aragón, Ruben Escribano, Reinaldo Rivera.

Resources: Ruben Escribano, Pamela Hidalgo.

Software: Manuela Pérez-Aragón, Reinaldo Rivera.

Supervision: Ruben Escribano.

Validation: Manuela Pérez-Aragón, Ruben Escribano, Reinaldo Rivera.

Visualization: Ruben Escribano, Reinaldo Rivera.

Writing – original draft: Manuela Pérez-Aragón, Ruben Escribano, Reinaldo Rivera.

Writing – review & editing: Ruben Escribano, Reinaldo Rivera, Pamela Hidalgo.

References

1. Angel MV. Biodiversity of the Pelagic Ocean. *Conserv Biol.* 1993; 7: 760–772.
2. Ormond RFG, Gage JD, Angel M V. *Marine Biodiversity: Patterns and Processes.* Cambridge University Press. 1997.
3. Currie DJ, Francis AP, Kerr JT. Some general propositions about the study of spatial patterns of species richness. *Ecoscience.* 1999; 6: 392–399. <https://doi.org/10.1080/11956860.1999.11682541>
4. Dornelas M, Gotelli NJ, McGill B, Shimadzu H, Moyes F, Sievers C, et al. Assemblage Time Series Reveal Biodiversity Change but Not Systematic Loss. *Science* (80-). 2014; 344: 296–300. <https://doi.org/10.1126/science.1248484> PMID: 24744374
5. Renema W, Bellwood DR, Braga JC, Bromfield K, Hall R, Johnson KG, et al. Hopping Hotspots: Global Shifts in Marine Biodiversity. *Science* (80-). 2008; 321: 654–657. <https://doi.org/10.1126/science.1155674> PMID: 18669854
6. Dornelas M, Antão LH, Moyes F, Et al. BioTIME: A database of biodiversity time series for the Anthropocene. *Glob Ecol Biogeogr.* 2018; 27: 760–786. <https://doi.org/10.1111/geb.12729> PMID: 30147447
7. Verity PG, Smetacek V. Organism life cycles, predation, and the structure of marine pelagic ecosystems. *Mar Ecol Prog Ser.* 1996; 130: 277–293. <https://doi.org/10.3354/meps130277>
8. Lombard F, Boss E, Waite AM, Vogt M, Uitz J, Stemmann L, et al. Globally Consistent Quantitative Observations of Planktonic Ecosystems. *Front Mar Sci.* 2019; 6. <https://doi.org/10.3389/fmars.2019.00196>
9. Valdés V, Carlotti F, Escribano R, Donoso K, Pagano M, Molina V, et al. Nitrogen and phosphorus recycling mediated by copepods and response of bacterioplankton community from three contrasting areas in the western tropical South Pacific (20° S). 2018; 6019–6032.
10. Escribano R, Hidalgo P, Valdés V, Frederick L. Temperature effects on development and reproduction of copepods in the Humboldt Current: the advantage of rapid growth. *J Plankton Res.* 2014; 36: 104–116. <https://doi.org/10.1093/plankt/fbt095>
11. Ruz PM, Hidalgo P, Escribano R, Keister JE, Yebra L, Franco-Cisterna B. Hypoxia effects on females and early stages of *Calanus chilensis* in the Humboldt Current ecosystem (23° S). *J Exp Mar Bio Ecol.* 2018; 498: 61–71. <https://doi.org/10.1016/j.jembe.2017.09.018>
12. Thor P, Dupont S. Transgenerational effects alleviate severe fecundity loss during ocean acidification in a ubiquitous planktonic copepod. *Glob Chang Biol.* 2015; 21: 2261–2271. <https://doi.org/10.1111/gcb.12815> PMID: 25430823
13. Lewis CN, Brown KA, Edwards LA, Cooper G, Findlay HS. Sensitivity to ocean acidification parallels natural pCO₂ gradients experienced by Arctic copepods under winter sea ice. *Proc Natl Acad Sci.* 2013; 110: E4960–E4967. <https://doi.org/10.1073/pnas.1315162110> PMID: 24297880
14. Peterson WT, Bellantoni DC. Relationships between water-column stratification, phytoplankton cell size and copepod fecundity in long island sound and off central Chile. *South African J Mar Sci.* 1987; 5: 411–421. <https://doi.org/10.2989/025776187784522748>
15. Williams R, Conway DVP, Hunt HG. The role of copepods in the planktonic ecosystems of mixed and stratified waters of the European shelf seas. *Hydrobiologia.* 1994; 292–293: 521–530. <https://doi.org/10.1007/BF00229980>
16. Kiørboe T, Nielsen TG. Regulation of zooplankton biomass and production in a temperate, coastal ecosystem. 1. Copepods. 1994; 39: 493–507. Available: <https://play.google.com/store/apps/details?id=com.evozi.deviceid&hl=pl>

17. Berline L, Siokou-Frangou I, Marasovic I, Vidjak O, Fernández de Puellas ML, Mazzocchi MG, et al. Intercomparison of six Mediterranean zooplankton time series. *Prog Oceanogr*. 2012; 97–100: 76–91. <https://doi.org/10.1016/j.pocean.2011.11.011>
18. Pino-Pinuer P, Escribano R, Hidalgo P, Riquelme-Bugueño R, Schneider W. Copepod community response to variable upwelling conditions off central-southern Chile during 2002–2004 and 2010–2012. *Mar Ecol Prog Ser*. 2014; 515: 83–95. <https://doi.org/10.3354/meps11001>
19. Medellín-Mora J, Escribano R, Schneider W. Community response of zooplankton to oceanographic changes (2002–2012) in the central/southern upwelling system of Chile. *Prog Oceanogr*. 2016; 142: 17–29. <https://doi.org/10.1016/j.pocean.2016.01.005>
20. Frederick L, Escribano R, Morales CE, Hormazabal S, Medellín-Mora J. Mesozooplankton respiration and community structure in a seamount region of the eastern South Pacific. *Deep Res Part I*. 2018; 135: 74–87. <https://doi.org/10.1016/j.dsr.2018.03.008>
21. Costello MJ, Chaudhary C. Marine Biodiversity, Biogeography, Deep-Sea Gradients, and Conservation. *Curr Biol*. 2017; 27: R511–R527. <https://doi.org/10.1016/j.cub.2017.04.060> PMID: 28586689
22. Ocean Biogeographic Information System (OBIS). Available: <http://iobis.org/>
23. Rombouts I, Beaugrand G, Ibañez F, Gasparini S, Chiba S, Legendre L. Global latitudinal variations in marine copepod diversity and environmental factors. *Proc R Soc B Biol Sci*. 2009; 276: 3053–3062. <https://doi.org/10.1098/rspb.2009.0742> PMID: 19515670
24. Hooff RC, Peterson WT. Copepod Biodiversity as an Indicator of Changes in Ocean and Climate Conditions of the Northern California Current Ecosystem. *Limnol Oceanogr*. 2006; 51: 2607–2620.
25. Brandão MC, Benedetti F, Martini S, Soviadan YD, Irsson JO, Romagnan JB, et al. Macroscale patterns of oceanic zooplankton composition and size structure. *Sci Rep*. 2021; 11: 1–19. <https://doi.org/10.1038/s41598-021-94615-5> PMID: 34344925
26. Muller-Karger FE, Canonico G, Aguilar CB, Bax NJ, Appeltans W, Yarincik K, et al. Marine Life 2030: building global knowledge of marine life for local action in the Ocean Decade. *ICES J Mar Sci*. 2023; 80: 355–357. <https://doi.org/10.1093/icesjms/fsac084>
27. Reid JL, Brinton E, Fleminger A, Venrick EL, McGowan JA. Ocean Circulation and Marine Life. *Adv Oceanogr*. 1978; 65–130. https://doi.org/10.1007/978-1-4615-8273-1_3
28. Wright DH. Species-energy theory: an extension of species-area theory. *Oikos*. 1983; 41: 496–506.
29. Currie DJ. Energy and large-scale patterns of animal and plant species richness. *Am Nat*. 1991; 137: 27–49. <https://doi.org/10.1086/285144>
30. Allen AP, Brown JH, Gillooly JF. Global biodiversity, biochemical kinetics, and the energetic-equivalence rule. *Science (80-)*. 2002; 297: 1545–1548. <https://doi.org/10.1126/science.1072380> PMID: 12202828
31. Rohde K. Latitudinal Gradients in Species Diversity: The Search for the Primary Cause. *Oikos*. 1992; 65: 514–527.
32. Roy K, Jablonski D, Valentine JW, Rosenberg G. Marine latitudinal diversity gradients: Tests of causal hypotheses. *Proc Natl Acad Sci*. 2002; 95: 3699–3702. <https://doi.org/10.1073/pnas.95.7.3699> PMID: 9520429
33. Brown JH, Gillooly JF, Allen AP, Savage VM, West GB. Toward a metabolic theory of ecology. *Ecol*. 2004; 85: 1771–1789.
34. Tittensor DP, Mora C, Jetz W, Lotze HK, Ricard D, Berghe E Vanden, et al. Global patterns and predictors of marine biodiversity across taxa. *Nature*. 2010; 466: 1098–1101. <https://doi.org/10.1038/nature09329> PMID: 20668450
35. Beaugrand G, Edwards M, Legendre L. Marine biodiversity, ecosystem functioning, and carbon cycles. *Proc Natl Acad Sci*. 2010; 107: 10120–10124. <https://doi.org/10.1073/pnas.0913855107> PMID: 20479247
36. Woodd-Walker RS, Ward P, Clarke A. Large-scale patterns in diversity and community structure of surface water copepods from the Atlantic Ocean. *Mar Ecol Prog Ser*. 2002; 236: 189–203. <https://doi.org/10.3354/meps236189>
37. Richardson AJ, Schoeman DS. Climate impact on plankton ecosystems in the Northeast Atlantic. *Science (80-)*. 2004; 305: 1609–1612. <https://doi.org/10.1126/science.1100958> PMID: 15361622
38. González CE, Medellín-Mora J, Escribano R. Environmental Gradients and Spatial Patterns of Calanoid Copepods in the Southeast Pacific. *Front Ecol Evol*. 2020; 8: 1–16. <https://doi.org/10.3389/fevo.2020.554409>
39. Suckow MA, Weisbroth SH, Franklin CL. Temperature in the oceans. *Seawater: Its Composition, Properties and Behaviour*. 1995. pp. 14–28. <https://doi.org/10.1016/B978-075063715-2/50003-4>

40. Ibarbalz FM, Henry N, Brandão MC, Martini S, Bussen G, Byrne H, et al. Global Trends in Marine Plankton Diversity across Kingdoms of Life. *Cell*. 2019; 179: 1084–1097.e21. <https://doi.org/10.1016/j.cell.2019.10.008> PMID: 31730851
41. Rombouts I, Beaugrand G, Ibañez F, Gasparini S, Chiba S, Legendre L. A multivariate approach to large-scale variation in marine planktonic copepod diversity and its environmental correlates. *Limnol Oceanogr*. 2010; 55: 2219–2229. <https://doi.org/10.4319/lo.2010.55.5.2219>
42. Costello MJ, Cheung A, De Hauwere N. Surface Area and the Seabed Area, Volume, Depth, Slope, and Topographic Variation for the World's Seas, Oceans, and Countries. *Environ Sci Technol*. 2010; 44: 8821–8828. <https://doi.org/10.1021/es1012752> PMID: 21033734
43. International Hydrographic Organization. Limits of Oceans and Seas, 3rd Edition. Monaco; 1953. pp. 1–42.
44. International Hydrographic Organization. South Pacific Ocean and its sub-divisions. 4th Editio. Limits of Oceans and Seas, 4th Edition. 4th Editio. Monaco; 2002. pp. 1–16. Available: https://www.iho.int/mtg_docs/com_wg/S-23WG/S-23WG_Misc/Draft_2002/Draft_2002.htm
45. Provoost P, Bosch S. robis: Ocean Biodiversity Information System (OBIS) Client. R package version 2.8.2. 2021. Available: <https://cran.r-project.org/package=robis>
46. Wickham H, Hester J, Chang W. devtools: Tools to Make Developing R Packages Easier. R package version 2.4.3. 2021. Available: <https://cran.r-project.org/package=devtools>
47. Desportes C, Drévillon M, Clavier M, Gounou A. Global Ocean Ensemble Physics Reanalysis—Low resolution. Available: https://data.marine.copernicus.eu/product/GLOBAL_REANALYSIS_PHY_001_026/description
48. Le Galloudec O, Perruche C, Derval C, Tressol M, Dussurget R. Global ocean biogeochemistry hind-cast. Available: https://data.marine.copernicus.eu/product/GLOBAL_MULTIYEAR_BGC_001_029/description
49. Guinehut S. Multi Observation Global Ocean 3D Temperature Salinity Height Geostrophic Current and MLD. Available: https://data.marine.copernicus.eu/product/MULTIOBS_GLO_PHY_TSUV_3D_MYNRT_015_012/description
50. Owens HL, Guralnick R. climateStability: An R package to estimate climate stability from time-slice climatologies. *Biodivers Informatics*. 2019; 14: 8–13. <https://doi.org/10.17161/bi.v14i0.9786>
51. Genuer R, Poggi J-M, Tuleau-Malot C. Variable Selection Using Random Forests. 2019. Available: <https://cran.r-project.org/web/packages/VSURF/VSURF.pdf>
52. Genuer R, Poggi J-M, Tuleau-Malot C. VSURF: An R Package for Variable Selection Using Random Forests. *R J*. 2015; 7: 19–33. Available: <https://journal.r-project.org/archive/2015-2/genuer-poggi-tuleaumalot.pdf>
53. Barbosa AM. fuzzySim: Applying fuzzy logic to binary similarity indices in ecology. *Methods Ecol Evol*. 2015; 6: 853–858. <https://doi.org/10.1111/2041-210X.12372>
54. Hammer Ø, Harper DAT, Ryan PD. Past: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontol Electron*. 2001; 4: 9pp. Available: http://palaeo-electronica.orghttp://palaeo-electronica.org/2001_1/past/issue1_01.htm.
55. Getis A, Ord JK. The Analysis of Spatial Association by Use of Distance Statistics. *Geogr Anal*. 1992; 24: 189–206. <https://doi.org/10.1111/j.1538-4632.1992.tb00261.x>
56. Swingland IR. Biodiversity, Definition of. *Encycl Biodivers*. 2013; 1: 377–391. <https://doi.org/10.1016/B978-0-12-384719-5.00009-5>
57. Koleff P, Gaston KJ, Lennon JJ. Measuring beta diversity for presence–absence data. *J Anim Ecol*. 2003; 72: 367–382. <https://doi.org/10.1007/s00228-017-2223-5>
58. Baselga A, Gómez-Rodríguez C. Alpha, beta and gamma diversity: measuring differences in biological communities. *Nov Acta Científica Compostel*. 2019; 26: 39–45.
59. Qian H, Ricklefs RE, White PS. Beta diversity of angiosperms in temperate floras of eastern Asia and eastern North America. *Ecol Lett*. 2005; 8: 15–22. <https://doi.org/10.1111/j.1461-0248.2004.00682.x>
60. Laffan SW, Lubarsky E, Rosauer DF. Biodiverse, a tool for the spatial analysis of biological and related diversity. *Ecography (Cop)*. 2010; 33: 643–647. <https://doi.org/10.1111/j.1600-0587.2010.06237.x>
61. Baselga A, Orme CDL. Betapart: An R package for the study of beta diversity. *Methods Ecol Evol*. 2012; 3: 808–812. <https://doi.org/10.1111/j.2041-210X.2012.00224.x>
62. Sanches Melo A. Package “CommEcol”, Community Ecology Analyses. 2021. pp. 1–34.
63. Vilela B, Villalobos F. LetsR: A new R package for data handling and analysis in macroecology. *Methods Ecol Evol*. 2015; 6: 1229–1234. <https://doi.org/10.1111/2041-210X.12401>
64. Laffan SW. Indices available in Biodiverse. 2022. Available: <https://github.com/shawnlaffan/biodiverse/wiki/Indices#simpson-and-shannon>

65. Hsieh TC, Ma KH, Chao A. iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods Ecol Evol.* 2016; 7: 1451–1456. <https://doi.org/10.1111/2041-210X.12613>
66. Hamilton S. Ocean Currents Data and Map. Harvard Dataverse, V1; 2018. <https://doi.org/10.7910/DVN/TKGO2Z>
67. Crase B, Liedloff AC, Wintle BA. A new method for dealing with residual spatial autocorrelation in species distribution models. *Ecography (Cop).* 2012; 35: 879–888. <https://doi.org/10.1111/j.1600-0587.2011.07138.x>
68. Bivand RS, Wong DWS. Comparing implementations of global and local indicators of spatial association. *Test.* 2018; 27: 716–748. <https://doi.org/10.1007/s11749-018-0599-x>
69. Bivand R. spdep: Spatial Dependence: Weighting Schemes, Statistics. 2023. p. 168. <https://doi.org/10.1111/gean.12319>
70. Fabozzi FJ, Focardi SM, Rachev ST, Arshanapalli BG. Model Selection Criterion: AIC and BIC. *The Basics of Financial Econometrics.* John Wiley & Sons, Inc.; 2014. pp. 399–403. <https://doi.org/10.1002/9781118856406.app5>
71. Bartoń K. Package MuMIn: Multi-Model Inference. 2022. Available: <https://cran.r-project.org/web/packages/MuMIn/>
72. Mokany K, Ware C, Woolley SNC, Ferrier S, Fitzpatrick MC. A working guide to harnessing generalized dissimilarity modelling for biodiversity analysis and conservation assessment. *Glob Ecol Biogeogr.* 2022; 31: 802–821. <https://doi.org/10.1111/geb.13459>
73. Fitzpatrick M, Mokany K, Manion G, Nieto-Lugilde D, FFerrier S, Lisk M, et al. Package “gdm.” 2022. <https://doi.org/10.1111/geb.13459>
74. Schneider W, Fukasawa M, Garcés-Vargas J, Bravo L, Uchida H, Kawano T, et al. Spin-up of South Pacific subtropical gyre freshens and cools the upper layer of the eastern South Pacific Ocean. *Geophys Res Lett.* 2007; 34: 1–5. <https://doi.org/10.1029/2007GL031933>
75. Reid J. On the total geostrophic circulation of the South Pacific Ocean: Flow patterns. *Tracers.* 1986; and transp: Vol. 16.
76. Qu T, Lindstrom EJ. A Climatological Interpretation of the Circulation in the Western South Pacific*. *J Phys Oceanogr.* 2011; 32: 2492–2508. <https://doi.org/10.1175/1520-0485-32.9.2492>
77. Ganachaud A, Cravatte S, Melet A, Schiller A, Holbrook NJ, Sloyan BM, et al. The Southwest Pacific circulation and climate experiment (SPICE). *J Geophys Res Ocean.* 2014; 119: 7660–7686. <https://doi.org/10.1002/2013JC009678>. Received
78. Ridgway KR, Dunn JR. Mesoscale structure of the mean East Australian Current System and its relationship with topography. *Prog Oceanogr.* 2003; 56: 189–222. [https://doi.org/10.1016/S0079-6611\(03\)00004-1](https://doi.org/10.1016/S0079-6611(03)00004-1)
79. Rintoul SR, Hughes CW, Olbers D. The Antarctic Circumpolar Current System. In: Siedler G, Church J, Gould J, editors. *Ocean Circulation and Climate: Observing and Modelling the Global Ocean.* 2001. pp. 271–302.
80. Stramma L, Peterson RG, Tomczak M. The South Pacific Current. *J Phys Oceanogr.* 1995; 25: 77–91. [https://doi.org/10.1175/1520-0485\(1995\)025<0077:tspc>2.0.co;2](https://doi.org/10.1175/1520-0485(1995)025<0077:tspc>2.0.co;2)
81. Montecino V, Lange CB. The Humboldt Current System: Ecosystem components and processes, fisheries, and sediment studies. *Prog Oceanogr.* 2009; 83: 65–79. <https://doi.org/10.1016/j.pocean.2009.07.041>
82. Merino-Campos V, De Pol-Holz R, Southon J, Latorre C, Collado-Fabbri S. Marine radiocarbon reservoir age along the Chilean continental margin. *Radiocarbon.* 2019; 61: 195–210. <https://doi.org/10.1017/RDC.2018.81>
83. García-Reyes M, Sydeman WJ, Schoeman DS, Rykaczewski RR, Black BA, Smit AJ, et al. Under Pressure: Climate Change, Upwelling, and Eastern Boundary Upwelling Ecosystems. *Front Mar Sci.* 2015; 2: 1–10. <https://doi.org/10.3389/fmars.2015.00109>
84. Strub PT, Mesías JM, Montecino V, Rutllant J, Salinas S. Coastal Ocean Circulation off Western South America. Volume 11. *The Sea.* Volume 11. John Wiley & Sons, Inc.; 1998. pp. 273–313.
85. Ancapichun, Santiago; Garcés-Vargas J. Variability of the Southeast Pacific Subtropical Anticyclone and its impact on sea surface temperature off north-central Chile. *Ciencias Mar.* 2015; 41: 1–20. <https://doi.org/10.7773/cm.v41i1.2338>
86. Carlotti F, Pagano M, Guilloux L, Donoso K, Valdés V, Grosso O, et al. Meso-zooplankton structure and functioning in the western tropical South Pacific along the 20th parallel south during the OUT-PAVE survey (February–April 2015). *Biogeosciences.* 2018; 15: 7273–7297. <https://doi.org/10.5194/bg-15-7273-2018>

87. Le Borgne R, Rodier M. Net zooplankton and the biological pump: A comparison between the oligotrophic and mesotrophic equatorial Pacific. *Deep Res Part II Top Stud Oceanogr.* 1997; 44: 2003–2023. [https://doi.org/10.1016/S0967-0645\(97\)00034-9](https://doi.org/10.1016/S0967-0645(97)00034-9)
88. Le Borgne R, Champalbert G, Gaudy R. Mesozooplankton biomass and composition in the equatorial Pacific along 180. *J Geophys Res Ocean.* 2003; 108: 1–10. <https://doi.org/10.1029/2000jc000745>
89. Dai L, Li C, Yang G, Sun X. Zooplankton abundance, biovolume and size spectra at western boundary currents in the subtropical North Pacific during winter 2012. *J Mar Syst.* 2016; 155: 73–83. <https://doi.org/10.1016/j.jmarsys.2015.11.004>
90. Fernández-Álamo MA, Färber-Lorda J. Zooplankton and the oceanography of the eastern tropical Pacific: A review. *Prog Oceanogr.* 2006; 69: 318–359. <https://doi.org/10.1016/j.pocean.2006.03.003>
91. Pennington JT, Mahoney KL, Kuwahara VS, Kolber DD, Calienes R, Chavez FP. Primary production in the eastern tropical Pacific: A review. *Prog Oceanogr.* 2006; 69: 285–317. <https://doi.org/10.1016/j.pocean.2006.03.012>
92. Dunstan PK, Hayes D, Woolley S, Allain V, Leduc D, Flynn A, et al. Bioregions of the South West Pacific Ocean. 2018. Available: https://rmi-data.sprep.org/system/files/Preliminary-worskop-report-Pacific_v2_1.pdf
93. Hidalgo P, Escribano R, Vergara O, Jorquera E, Donoso K, Mendoza P. Patterns of copepod diversity in the Chilean coastal upwelling system. *Deep Res Part II.* 2010; 57: 2089–2097. <https://doi.org/10.1016/j.dsr2.2010.09.012>
94. Dolan JR, Lemée R, Gasparini S, Mousseau L, Heyndrickx C. Probing diversity in the plankton: using patterns in Tintinnids (planktonic marine ciliates) to identify mechanisms. *Hydrobiologia.* 2006; 555: 143–157. <https://doi.org/10.1007/s10750-005-1112-6>
95. Rutherford S, Hondt SD, Prell W. Environmental controls on the geographic distribution of zooplankton diversity. *Nature.* 1999; 400: 749–753.
96. Worm B, Sandow M, Oschlies A, Lotze HK, Myers RA. Global patterns of predator diversity in the open oceans. *Science (80-).* 2005; 309: 1365–1369. <https://doi.org/10.1126/science.1113399> PMID: 16051749
97. McGowan JA, Walker PW. Structure in the Copepod Community of the North Pacific Central Gyre. *Ecol Monogr.* 1979; 49: 195–226. Available: <http://www.jstor.org/stable/1942513>
98. Vereshchaka A, Abyzova G, Lunina A, Musaeva E. The deep-sea zooplankton of the North, Central, and South Atlantic: Biomass, abundance, diversity. *Deep Res Part II Top Stud Oceanogr.* 2017; 137: 89–101. <https://doi.org/10.1016/j.dsr2.2016.06.017>
99. Rivera R, Escribano R, González CE, Pérez-Aragón M. Modeling present and future distribution of plankton populations in a coastal upwelling zone: the copepod *Calanus chilensis* as a study case. *Sci Rep.* 2023; 13. <https://doi.org/10.1038/s41598-023-29541-9> PMID: 36823290
100. Rigby S, Milsom C. Benthic origins of zooplankton: An environmentally determined macroevolutionary effect. *Geology.* 1996; 24: 52–54. [https://doi.org/10.1130/0091-7613\(1996\)024<0052:BOOZAE>2.3.CO;2](https://doi.org/10.1130/0091-7613(1996)024<0052:BOOZAE>2.3.CO;2)
101. Bradford-Grieve JM. Colonization of the pelagic realm by calanoid copepods. *Hydrobiologia.* 2002; 485: 223–244. <https://doi.org/10.1023/A:1021373412738>
102. Braga E, Zardoya R, Meyer A, Yen J. Mitochondrial and nuclear rRNA based copepod phylogeny with emphasis on the Euchaetidae (Calanoida). *Mar Biol.* 1999; 133: 79–90. <https://doi.org/10.1007/s002270050445>
103. Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G, et al. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography (Cop).* 2013; 36: 27–46. <https://doi.org/10.1111/j.1600-0587.2012.07348.x>
104. Sverdrup HU, Johnson MW, Fleming RH. Dissolved Gases in Sea Water. Chapter VI: Chemistry of Sea Water. The Oceans: Their Physics, Chemistry, and General Biology. New York: Prentice-Hall, Inc.; 1942. pp. 186–191. Available: <https://publishing.cdlib.org/ucpressebooks/view?docId=kt167nb66r;brand=eschol>
105. White JR, Zhang X, Welling LA, Roman MR, Dam HG. Latitudinal gradients in zooplankton biomass in the tropical Pacific at 140°W during the JGOFS EqPac study: Effects of El Niño. *Deep Res Part II.* 1995; 42: 715–733. [https://doi.org/10.1016/0967-0645\(95\)00033-M](https://doi.org/10.1016/0967-0645(95)00033-M)
106. Forster J, Hirst AG, Atkinson D. Warming-induced reductions in body size are greater in aquatic than terrestrial species. *Proc Natl Acad Sci U S A.* 2012; 109: 19310–19314. <https://doi.org/10.1073/pnas.1210460109> PMID: 23129645
107. Roman MR, Pierson JJ. Interactive Effects of Increasing Temperature and Decreasing Oxygen on Coastal Copepods. *Biol Bull.* 2022; 000–000. <https://doi.org/10.1086/722111> PMID: 36548979

108. Wishner KF, Seibel B, Outram D. Ocean deoxygenation and copepods: Coping with oxygen minimum zone variability. *Biogeosciences*. 2020; 17: 2315–2339. <https://doi.org/10.5194/bg-17-2315-2020>
109. Paffenhöfer G-A. Oxygen consumption in relation to motion of marine planktonic copepods. *Mar Ecol Prog Ser*. 2006; 317: 187–192.
110. Cervetto G, Gaudy R, Pagano M. Influence of salinity on the distribution of *Acartia tonsa* (Copepoda, Calanoida). *J Exp Mar Bio Ecol*. 1999.
111. Vuorinen I, Hänninen J, Viitasalo M, Helminen U, Kuosa H. Proportion of copepod biomass declines with decreasing salinity in the Baltic Sea. *ICES J Mar Sci*. 1998.
112. Choi SY, Lee EH, Soh HY, Jang MC. Effects of Temperature and Salinity on Egg Production, Hatching, and Mortality Rates in *Acartia ohtsukai* (Copepoda, Calanoida). *Front Mar Sci*. 2021; 8. <https://doi.org/10.3389/fmars.2021.704479>
113. Dutz J, Christensen AM. Broad plasticity in the salinity tolerance of a marine copepod species, *Acartia longiremis*, in the Baltic Sea. *J Plankton Res*. 2018; 40: 342–355. <https://doi.org/10.1093/plankt/fby013>
114. Milione M, Zeng C. The effects of temperature and salinity on population growth and egg hatching success of the tropical calanoid copepod, *Acartia sinjiensis*. *Aquaculture*. 2008; 275: 116–123. <https://doi.org/10.1016/j.aquaculture.2007.12.010>
115. Lee SH, Lee MC, Puthumana J, Park JC, Kang S, Hwang DS, et al. Effects of salinity on growth, fatty acid synthesis, and expression of stress response genes in the cyclopoid copepod *Paracyclops nana*. *Aquaculture*. 2017; 470: 182–189. <https://doi.org/10.1016/j.aquaculture.2016.12.037>
116. Pörtner HO. Climate variations and the physiological basis of temperature dependent biogeography: Systemic to molecular hierarchy of thermal tolerance in animals. *Comp Biochem Physiol—A Mol Integr Physiol*. 2002; 132: 739–761. [https://doi.org/10.1016/s1095-6433\(02\)00045-4](https://doi.org/10.1016/s1095-6433(02)00045-4) PMID: 12095860
117. Elliott DT, Tang KW. Influence of carcass abundance on estimates of mortality and assessment of population dynamics in *acartia tonsa*. *Mar Ecol Prog Ser*. 2011; 427: 1–12. <https://doi.org/10.3354/meps09063>
118. Sorte CJB, Jones SJ, Miller LP. Geographic variation in temperature tolerance as an indicator of potential population responses to climate change. *J Exp Mar Bio Ecol*. 2011; 400: 209–217. <https://doi.org/10.1016/j.jembe.2011.02.009>
119. Mackas DL, Batten S, Trudel M. Effects on zooplankton of a warmer ocean: Recent evidence from the Northeast Pacific. *Prog Oceanogr*. 2007; 75: 223–252. <https://doi.org/10.1016/j.pocean.2007.08.010>
120. Barton AD, Pershing AJ, Litchman E, Record NR, Edwards KF, Finkel Z V., et al. The biogeography of marine plankton traits. *Ecol Lett*. 2013; 16: 522–534. <https://doi.org/10.1111/ele.12063> PMID: 23360597
121. Benedetti F, Wydler J, Vogt M. Copepod functional traits and groups show contrasting biogeographies in the global ocean. *BioRxiv*. 2022. <https://doi.org/10.1101/2022.02.24.481747>
122. González CE, Goetze E, Escrí R, Ulloa O, Victoriano P. Genetic diversity and novel lineages in the cosmopolitan copepod *Pleuromamma abdominalis* in the Southeast Pacific. *Sci Rep*. 2020; 1–15. <https://doi.org/10.1038/s41598-019-56935-5> PMID: 31980660
123. Turner JT. Latitudinal Patterns of Calanoid and Cyclopoid Copepod Diversity in Estuarine Waters of Eastern North America. *J Biogeogr*. 1981; 8: 369–382. <https://doi.org/10.2307/2844757>
124. Cornils A, Blanco-Bercial L. Phylogeny of the Paracalanidae Giesbrecht, 1888 (Crustacea: Copepoda: Calanoida). *Mol Phylogenet Evol*. 2013; 69: 861–872. <https://doi.org/10.1016/j.ympev.2013.06.018> PMID: 23831457
125. Escríbano R, Bustos-ríos E, Hidalgo P, Morales CE. Non-limiting food conditions for growth and production of the copepod community in a highly productive upwelling zone. *Cont Shelf Res*. 2016; 126: 1–14. <https://doi.org/10.1016/j.csr.2016.07.018>
126. Brun P, Payne MR, Kiørboe T. A trait database for marine copepods. *Earth Syst Sci Data*. 2017; 9: 99–113. <https://doi.org/10.5194/essd-9-99-2017>
127. Medellín-Mora J, Escríbano R, Corredor-Acosta A, Hidalgo P, Schneider W. Uncovering the Composition and Diversity of Pelagic Copepods in the Oligotrophic Blue Water of the South Pacific Subtropical Gyre. *Front Mar Sci*. 2021; 8: 1–18. <https://doi.org/10.3389/fmars.2021.625842>
128. Razouls C, Desreumaux N, Kouwenberg J, de Bovée F. Biodiversity of Marine Planktonic Copepods (morphology, geographical distribution and biological data), 2005–2023. In: Sorbonne University, CNRS. [Internet]. 2023 [cited 25 Jan 2024]. Available: <http://copepodes.obs-banyuls.fr/en>
129. Boakes EH, McGowan PJK, Fuller RA, Chang-Qing D, Clark NE, O'Connor K, et al. Distorted views of biodiversity: Spatial and temporal bias in species occurrence data. *PLoS Biol*. 2010; 8. <https://doi.org/10.1371/journal.pbio.1000385> PMID: 20532234

130. Hortal J, De Bello F, Diniz-Filho JAF, Lewinsohn TM, Lobo JM, Ladle RJ. Seven Shortfalls that Beset Large-Scale Knowledge of Biodiversity. *Annu Rev Ecol Evol Syst*. 2015; 46: 523–549. <https://doi.org/10.1146/annurev-ecolsys-112414-054400>
131. Ball-Damerow JE, Brenskelle L, Barve N, Soltis PS, Sierwald P, Bieler R, et al. Research applications of primary biodiversity databases in the digital age. *PLoS One*. 2019; 14. <https://doi.org/10.1371/journal.pone.0215794> PMID: 31509534
132. Nelson G, Ellis S. The history and impact of digitization and digital data mobilization on biodiversity research. *Philosophical Transactions of the Royal Society B: Biological Sciences*. Royal Society Publishing; 2019. <https://doi.org/10.1098/rstb.2017.0391> PMID: 30455209
133. Page LM, Macfadden BJ, Fortes JA, Soltis PS, Riccardi G. Digitization of Biodiversity Collections Reveals Biggest Data on Biodiversity. *BioScience*. Oxford University Press; 2015. pp. 841–842. <https://doi.org/10.1093/biosci/biv104>
134. Ariño AH. Approaches to estimating the universe of natural history collections data. *Biodivers Informatics*. 2010; 7: 81–92. <https://doi.org/10.17161/bi.v7i2.3991>
135. Isaac NJB, van Strien AJ, August TA, de Zeeuw MP, Roy DB. Statistics for citizen science: Extracting signals of change from noisy ecological data. *Methods Ecol Evol*. 2014; 5: 1052–1060. <https://doi.org/10.1111/2041-210X.12254>
136. García-Roselló E, Guisande C, Manjarrés-Hernández A, González-Dacosta J, Heine J, Pelayo-Villamil P, et al. Can we derive macroecological patterns from primary Global Biodiversity Information Facility data? *Glob Ecol Biogeogr*. 2015; 24: 335–347. <https://doi.org/10.1111/geb.12260>
137. Gaylord B, Gaines SD. Temperature or transport? Range limits in marine species mediated solely by flow. *Am Nat*. 2000; 155: 769–789. <https://doi.org/10.1086/303357>
138. Navarrete SA, Broitman B, Wieters EA, Finke GR, Venegas RM, Sotomayor A. Recruitment of intertidal invertebrates in the southeast Pacific: Interannual variability and the 1997–1998 El Niño. *Limnol Ocean*. 2002; 47: 791–802. Available: <http://www.pfeg.noaa.gov>
139. Watson JR, Hays CG, Raimondi PT, Mitarai S, Dong C, McWilliams JC, et al. Currents connecting communities: nearshore community similarity and ocean circulation. *Ecology*. 2011. <https://doi.org/10.1890/10-1436.1> PMID: 21797147
140. Coleman MA, Feng M, Roughan M, Cetina-Heredia P, Connell SD. Temperate shelf water dispersal by Australian boundary currents: Implications for population connectivity. *Limnol Oceanogr Fluids Environ*. 2013; 3: 295–309. <https://doi.org/10.1215/21573689-2409306>
141. Wernberg T, Thomsen MS, Connell SD, Russell BD, Waters JM, Zuccarello GC, et al. The footprint of continental-scale ocean currents on the biogeography of seaweeds. *PLoS One*. 2013; 8. <https://doi.org/10.1371/journal.pone.0080168> PMID: 24260352