Resolution changes relationships: Optimizing sampling design using small scale zooplankton data

James Scott [a,b,*], Sophie Pitois [b], Veronique Creach [b], Gill Malin [a], Phil Culverhouse [c], Julian Tilbury [c]

[a] Centre for Ocean and Atmospheric Sciences, School of Environmental Sciences, University of East Anglia, Norwich, United Kingdom
[b] The Centre for Environment, Fisheries and Aquacultural Science, Lowestoft, United Kingdom
[c] Plankton Analytics Ltd., Plymouth, United Kingdom

ARTICLE INFO

ABSTRACT

Marine research surveys are an integral tool in understanding the marine environment. Recent technological advances have allowed the development of automated or semi-automated methods for the collection of marine data. These devices are often easily implemented on existing surveys and can collect data at finer spatiotemporal resolutions than traditional devices. We used two automated instruments: the Plankton Imager and FerryBox, to collect information on zooplankton, temperature, salinity and chlorophyll in the Celtic Sea. The resulting data were spatiotemporally aligned and merged to decreasing spatial resolutions to explore how distribution patterns and the relationship between variables change across different spatial resolutions. Relative standard deviation was used to describe variability of merged data within grid cells. All variables displayed large, area-wide spatial patterns excluding copepod size which remained consistent across the study area. Copepod biomass and abundance displayed high variations across small spatial scales. Decreasing the sampling resolution changed the description of the data where small spatial changes (those that occur over scales < 3 km) were lost and area wide patterns were emphasized. Furthermore, we found that the choice of resolution can affect both the statistical strength and significance of relationships with high variability at lower resolutions due to the mismatch between the scales of ecological processes and sampling. Determining the optimum sampling resolution to answer a specific question will be dependent upon several factors, mainly the variable measured, season, location and scale of process, which all drive variation. These considerations should be a key element of survey design, helping move towards an integrated approach for an improved understanding of ecosystem processes and gaining a more holistic description of the marine environment.

1. Introduction

Research surveys are fundamental in furthering our understanding of the marine environment. Motivated by providing a holistic, ecosystem approach to monitoring (Kupschus et al., 2016) or mandated by policy (European Commission, 2008; Danovaro et al., 2016), technological developments are helping surveys move toward increasingly interdisciplinary approaches (ICES, 2015; Doray et al., 2018). Installing automated technologies, which allow for continuous data collection with little human input, are a straightforward step in achieving this goal. Devices such as the FerryBox, used here, collects both physical and biological variables continuously and reports high frequency data throughout a survey (Petersen and Colijn, 2017). Due to their

continuous, automated nature, these data are readily available for (near) real time analysis, or retrospective, post-collection, analysis. These often easy to implement devices can reduce vessel costs (time, fuel or labor) when compared to traditional methods (deployment of nets) or towed imaging devices (e.g., deployment and recovery or reduced vessel speed while towing) and allow for an increased number of variables collected at no or little extra cost. Their use can help surveys stay within financial limitations (Bean et al., 2017; Pitois et al., 2018) and allow for optimized survey design (Kupschus et al., 2016) by easily increasing sampling coverage and intensity (Owens, 2014; Doray et al., 2018). These devices do not require an onboard expert, freeing up vessel space and further reducing costs. They can typically sample in all weather conditions, allowing for data collection in hard to sample locations or reduce the

---

time spent waiting for safe sampling conditions.

In recent decades, automated technologies have become commonplace in multiple marine disciplines. The FerryBox is one of many established options for continuous, automated sampling of physical parameters. Acoustic devices are used globally, commercially and scientifically, for fishing (Mann et al., 2008; Simmonds and MacLennan, 2008), marine mammal research (Johnson and Tyack, 2003; Johnson et al., 2009) and bathymetry (de Moustier, 1986). Other purpose-built devices sample a single component, for example, fish eggs (Checkley et al., 2000) or phytoplankton (Olson et al., 2018). The continuous sampling of zooplankton, globally important in carbon cycles (Steinberg et al., 2002; Steinberg and Landry, 2017), fisheries science (Beaugrand et al., 2003; Heath, 2005; Lauria et al., 2013) and used as climate change indicators (Taylor et al., 2002), provides a unique technological challenge arising from the difficulties with sampling the entire zooplankton component accurately. Zooplankton includes a wide range of sizes and behaviors and undergoes rapid, temporal and spatial changes, known as plankton patchiness (Mackas et al., 1985; Abraham, 1998), making replicable sampling difficult. These fine scales changes can be seen with traditional net haul data, which provide a 'snapshot' of the zooplankton but have very high replicate tow variability (Wiebe et al., 1968; Lee and McAlice, 1979; Skjoldal et al., 2013). Capture by netting, and analysis by microscopy, form the gold standard of zooplankton sampling and are principally responsible for our understanding of zooplankton ecology. Their continued use to maintain time series and use as a reliable method is essential to further understanding the zooplankton. Like all sampling devices, nets do suffer some limitations. Deploying plankton nets is time consuming and the collected sample is often preserved using hazardous chemicals for analysis on shore. These challenges have placed pressure on developing cost-effective methods (Danovaro et al., 2016) and in response, technological developments have resulted in a variety of newer devices (Wiebe and Benfield, 2003; Lombard et al., 2019).

The Plankton Imager (PI) is a continuous, automated, imaging device used to sample zooplankton. The PI takes images of all passing particles in seawater pumped onboard a ship (Culverhouse et al., 2016; Pitois et al., 2018). An initial study evaluating the first generation of the instrument (previously known as Plankton Image Analyzer, Pitois et al. 2018) against traditional net sampling, found good agreement in the spatial distribution of zooplankton abundances, although noted a portion of fragile organisms (e.g., Appendicularia) were likely to be damaged by the system pump and consequently under-sampled. The study also described the overall lower capture efficiency of the PI with discrepancies mainly resulting from image quality, such as blurred images, which made accurate classification challenging. In response, hardware changes have resolved these issues, resulting in much improved image quality. The PI has since been used to describe temporal changes in the mesozooplankton community (Scott et al., 2021). This study found that those fragile species (e.g., Appendicularia) are sampled in sufficient quantity to detect seasonal difference. More recently the application of the PI to ecological indicators has been tested (Pitois et al., 2021). To date, all published studies have used the PI for point sampling, similar to a deployed ring net, as opposed to continuous sampling. Here we used a new data extraction method to best take advantage of the PI's continuous nature.

The PI has been used alongside the FerryBox routinely during UK fisheries surveys in the Celtic Sea. We use data collected in parallel from these devices to explore small scale changes in the zooplankton in the context of physical parameters and the relationships therein. As automated devices and the ability to collect vast quantities of data become increasingly common place, a new challenge has emerged in that the data collection rate has become faster than the processing rate, resulting in data bottlenecks. It is therefore important to focus collection efforts to gather the correct type of information, at the required locations, times and scales to answer a particular question, balancing research needs with budget limitations. Here, we aim to explore how best to determine the optimal resolution appropriate for the target process or relationship to avoid mismatching between sampling resolution and ecological scales. These can be used to inform future survey design leading to an increasingly holistic survey description.

2. Materials and methods

All data were collected in the Celtic Sea from the 3rd of October to the 7th of November 2020 aboard the RV Cefas Endeavour as part of the PELTIC survey (PELagic ecosystems in the Western English Channel and eastern celTIC Seas) (ICES, 2015) (Fig. 1). All in situ data were collected using the ship's continuous flow system sampling at 4 m below sea level. Zooplankton data were collected using the PI (Pitois et al., 2018). Temperature, salinity and fluorescence were collected using the FerryBox (4H-JENA, Germany). Zooplankton data were sampled at night for consistency and to reduce the effect of vertical migration (Lampert, 1989; Pitois et al., 2018).

2.1. Plankton Imager (PI)

The PI was connected to the ship's continuous flow pump, sampling at 22 L min^{-1} with negligible downtime (Fig. 2). The inlet pipe and internal ships piping have various internal diameters larger than the flow cell which has an internal depth of 12.8 mm, giving a field of view of 10 μm × 20.48 mm. As sea water passes through the flow cell where all passing particles are imaged by a Basler 2048-70kc line scan camera with a scanning rate of 70,000 lines per second. Lines are then stitched together, and regions of interest (ROI) are extracted and saved as images. GPS, time and particle size data (area, length and width) are saved in the metadata of each image. The PI worked continuously throughout the survey. The PI has adjustable minimum and maximum size parameters (min. 100 μm to max. 2 cm). When using this range, the processing rate of the images could not keep up with their collection rate (i.e., the images are captured faster than they can be written to disk. For the survey, the size range was set from 180 μm to 2 cm. This reduced the number of captured images and allowed for a more manageable dataset for archiving, processing and analysis. Even with this reduced size range, a 1 month survey typically collects 1 tb of data.

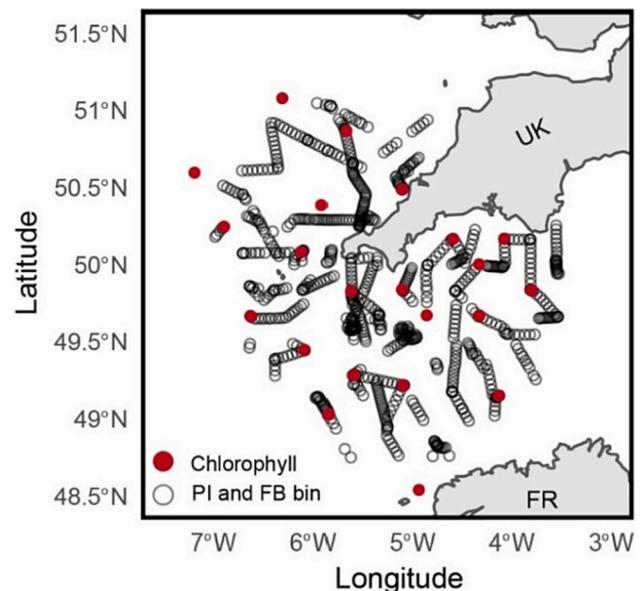Over 70 million images were collected during the survey. In the

Fig. 1. Celtic Sea study area and spatial extent of the collected data. Red filled symbols represent in situ discrete chlorophyll samples. Black open symbols represent PI and FerryBox (temperatures, salinity, fluorescence) 10-minute bins. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
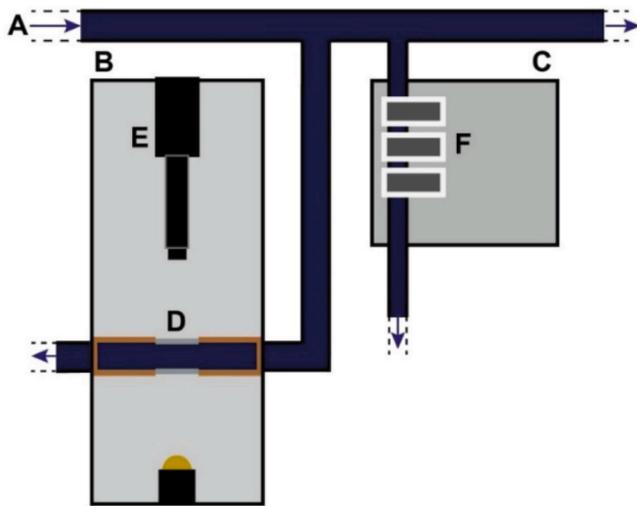
Fig. 2. Schematic of Plankton Imager (PI) and FerryBox setup aboard the RV Cefas Endeavor. Water is pumped onboard from 4 m below sea level (**A**). This supplies the PI (**B**) and FerryBox (**C**). Within the PI water flows through a flow cell (**D**) where passing particles are imaged by a line scan camera (**E**). Within the FerryBox water passes through a suite of sensors (**F**), here temperature, salinity and fluorescence are used.

absence of an accurate classifier for the PI, all images required manual classification. A series of subsets were used to reduce the number of images classified to an achievable quantity. A 0.25° grid was transposed over the study area. Each grid cell typically had multiple transects passing through with the specific number of transects varying based on vessel movements. Data were extracted from the shortest nighttime transect within each 0.25° cell (min = 20 mins, mean = 136 mins, max = 420 mins). The transect time (and therefore water sampled) varied within each grid cell dependent on vessel activities (e.g., steaming between stations or fishing). This extraction resulted in 17 million images for classification. Finally, data were temporally subsampled where 1 in 10 images were extracted from each transect to further reduce the size of the dataset. This process is similar to random subsampling or 'splitting' of a physical sample and assumes the distribution of organisms within a subsample follows a Poisson distribution (Postel et al., 2000). The resultant 1.7 million images were manually classified to "copepod" or "other" with copepod including the adult and copepodite stages and the latter category comprising all non-copepod zooplankton and detritus. Sorting to only these categories greatly sped up the classification process. The final copepod count per grid cell was multiplied by 10 to resolve for subsampling.

For statistical analysis, the selected transects were divided into 10 min bins where each bin sampled 0.22 m^3 of seawater. This totaled 853 bins (Fig. 1). The minimum bin size was determined as a compromise between obtaining the smallest possible spatial resolution while sampling a sufficient amount of water to allow for subsampling. Sampling a smaller amount of water (e.g. 1 min and 0.022 m^3 of water) may have resulted in unrealistic values when resolving for subsampling. Copepod density was reported as individuals per m^3 (indv. m^{-3}). Particle lengths were obtained from image metadata files and used as a proxy for copepod size or total length. Within each 10 min bin, the geometric mean (geomean) size of all individuals was calculated to take into account their non-normal distribution. This mean value was used to calculate mean copepod wet weight (i.e., individual biomass) following the equation from Pitois et al. 2021:

$$\text{copepod wet weight} = 0.299 \times \text{total length}^{2.8348}$$

This was then upscaled with copepod density to calculate biomass across the bin reported as mg m^{-3}.

2.2. In situ chlorophyll measurements

The FerryBox consists of a water inlet connected to the ship's continuous flow (Fig. 2). It comprises a suite of sensors for measuring physical variables (e.g., temperature, salinity, turbidity, fluorescence and oxygen) and corresponding metadata (GPS, date and time). All data are automatically bin averaged to 1 min on collection to save storage space. Only temperature (°C), salinity (psu) and fluorescence were used in this study. Discrete chlorophyll samples were taken from the continuous flow passing through the FerryBox at 22 locations within the study area (Fig. 1). The chlorophyll was extracted using 90 % acetone and measured with a Turner fluorometer (Strickland and Parsons, 1972). A linear model was used to convert the FerryBox chlorophyll fluorescence to chlorophyll. The fitted regression model was: [chlorophyll mg m^{-3} = 0.72 * chlorophyll fluorescence + 0.0637]. The regression was statistically significant (R^2 = 0.91, F = 206.1, p < 0.001). FerryBox data were spatiotemporally aligned to the 853, 10-minute copepod bins described above. This was achieved by taking the mean of three variables across the bin.

2.3. Analysis

At the finest resolution (the 853, 10 min bins), data were plotted to describe the broader spatial patterns and examine small spatial scale changes in all variables. For copepod biomass, size and density, the change in value between a 10 min bin and the previous 10 min bin was examined to see if there was a relationship between distance or time between bins and change in value. For statistical analysis and to investigate how changes in resolution can affect how spatial patterns are described and if small scale changes are omitted or accentuated, the 853 bins were merged to decreasing resolutions. Merging was achieved by taking the mean value of all merged bins for each variable. To explain variation within each cell at each resolution, Relative Standard Deviation (RSD) was used as it expresses the variability of a data set as a percentage relative to its location. RSD is calculated as: RSD = (sample standard deviation / sample mean) × 100.

Four resolutions, 0.1°, 0.25°, 0.5° and 1° were selected. The largest resolution was chosen based on the spatial extents of our study area. A resolution lower than 1° would have resulted in too few cells or a cell that contained the entire data. The selected resolutions were used to visually compare the changes in the description of spatial patterns associated with merging data to coarser resolutions. The relationship between RSD and decreasing resolution was also explored for all variables at resolutions between 0.01° and 0.9° decreasing in 0.01° increments. Here the mean RSD value across all cells was used. For statistical analyses, the same resolution range (0.01° to 0.9°, by = 0.01°) was used. Spearman's ρ coefficient was used to test for a significant relationship between number of stations per cell and RSD and explore the relationship between copepod biomass and chlorophyll at these resolutions. Copepod biomass, size and density were log-transformed (log10(x + 1)) for Figs. 3, 4 and 5 to highlight variability.

3. Results

3.1. Spatial distribution

Spatial patterns across the study area for copepod density and biomass were closely aligned (Fig. 3A, 3B). Higher copepod densities (>8000 indv. m^{-3}) and biomass (>150 mg m^{-3}) were typically found in the middle of the study area with lower values found toward the south (<2000 indv. m^{-3} and < 50 mg m^{-3}, respectively) (Fig. 3A, 3B). Density ranged from 45 to 8790 indv. m^{-3} and biomass ranged from < 1 to 155 mg m^{-3}. Copepod size had a more uniform distribution across the study area with no obvious spatial patterns with some localized exceptions of larger copepods found in the northern most extents of the study area (Fig. 3C). Size ranged from 199 to 2590 μm. Large fluctuations in each
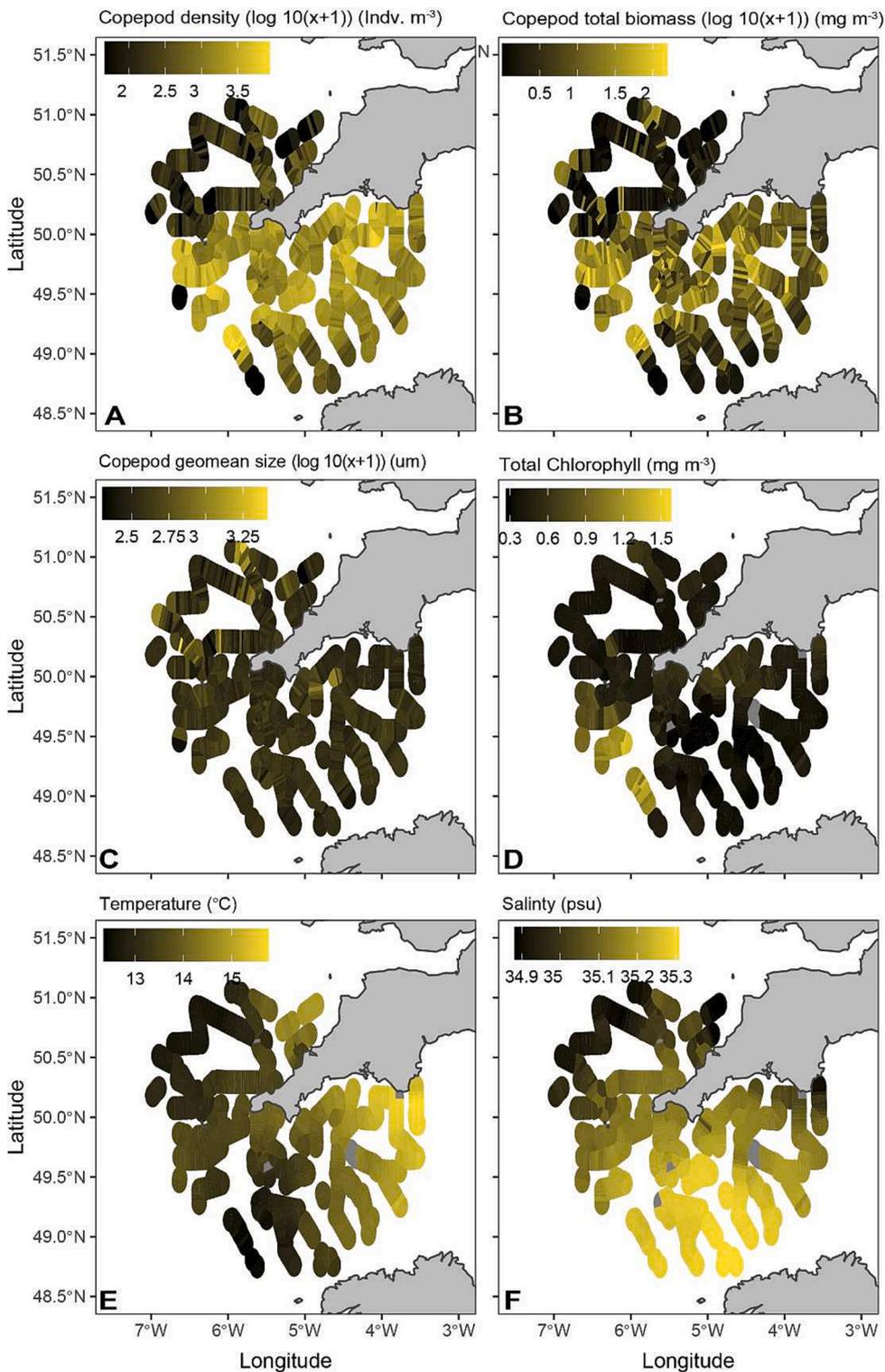
Fig. 3. Overview of each variable at the finest spatial resolution (10 min bins, approx. 2.2 m^{-3} seawater) as point data where each point is the bin median latitude and longitude. Point data are highlighted by using Voronoi triangles (with a maximum radius size around the point of 0.1°) which allows for a bigger point size, while avoiding overlap to better highlight small scale changes in the variable. For: (**A**) Copepod density (log10(x + 1)) (indv. m^{-3}), (**B**) Copepod total biomass (log10(x + 1)) (mg/m^{-3}(−|-)), (**C**) Copepod geomean size (μm), (**D**) Chlorophyll (mg/m^{-3}(−|-)), (**E**) Temperature (°C) and (**F**) Salinity (psu). Color scales are consistent with Fig. 5 and Fig. 8 for comparison.

variable were seen over small spatial scales (between adjacent bins, 5 km), this was less frequent for copepod size and is most evident in the central study area for copepod biomass (Fig. 3A – 3C).

Small scale changes were not present in chlorophyll concentration, temperature or salinity (Fig. 3D − 3F) with these variables displaying more gradual changes across the area. Temperature was higher toward the east of the study area (Fig. 3E) and salinity was higher toward the south (Fig. 3F). Chlorophyll concentration was consistently low (<0.6 mg m^{-3}) except for the most south-westerly extents of the study area

where the maximum value of 1.5 mg m^{-3} was seen (Fig. 3D).

The large variations across small spatial changes in all copepod variables are better highlighted by Fig. 4. The change in value between a 10-minute bin and the previous 10-minute bin was explored for density (Fig. 4A), size (Fig. 4B) and biomass (Fig. 4C). There was no clear relationship between the range and distance from the previous station for any variable at small spatial scales (<5 km). On the contrary, the highest changes in density and biomass tended to be within 3 km of the previous bin (Fig. 3A, B). This was seen most clearly in adjacent
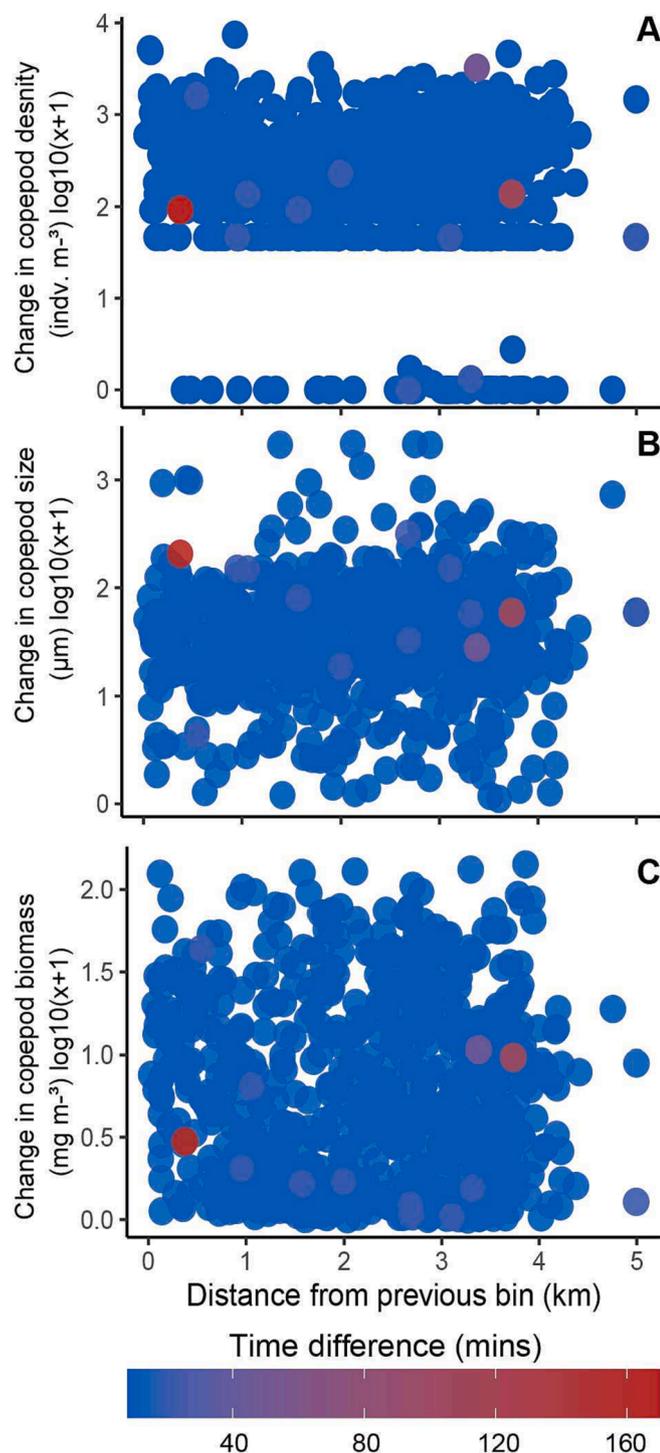
Fig. 4. Change in parameters value (y-axis) and time (color scales) between two bins as a function of their distance from each other (x-axis) for (**A**) copepod density (indv. m^{-3}), (**B**) copepod geomean size (μm) and (**C**) biomass (mg/m^{-3}(−|-)).

datapoints located in the middle of the study area (Fig. 3B, C).

3.2. Description at changing resolutions

Most resolutions captured the broad spatial patterns evident in the smallest resolution for copepod biomass, density and size (Fig. 5, Fig. 3A - C). For example, the regions of low biomass toward the north and higher biomass toward the southwest of the study area (Fig. 3B) were visible at all resolutions (Fig. 5). Although, large changes in copepod variables over small spatial scales seen at the smallest resolution were partly lost at a 0.1° resolution and absent entirely at 1° (Fig. 5). This is true for all copepod variables where a high level of detail was lost by only halving the resolution. For example, the area of low biomass (7° W, 49° N) seen at 0.5° resolution was lost when halving to 1° (Fig. 5, column 1). This loss of small scale detail while capturing broad patterns with decreased resolution was mirrored by copepod density and size. RSD is shown spatially for the selected.

resolutions (Fig. 5) and in increasing 0.01° increments in a scatter plot (Fig. 6) for all copepod variables. Using Spearmans ρ, copepod mean density RSD and number of datapoints per cell were consistently significantly related in cells < 0.27° resolution (at 0.26°, Rs = 0.31, p < 0.05, n = 78). For mean biomass RSD, there was consistent significance for cells < 0.4° (at 0.39°, Rs = 0.29, p = 0.05, n = 44). For mean size RSD there was no consistent significant relationship at any resolution. For all three-copepod variables, there were exceptions seen at lower resolutions which may result from an insufficient sample size for the Spearmans ρ test. Copepod sizes were consistently low in RSD (<30 %) between grid cells both spatially and across resolutions (Fig. 5, column 3; Supplementary Fig. 1). There was an increase in mean RSD, from 9.96 % to 30.55 %, with decreasing resolution (Fig. 6; Supplementary Fig. 1), although marginal when compared to other variables. Biomass had the highest spatial variation in RSD at all resolutions (Fig. 5, column 1; Supplementary Fig. 1). There was a larger increase meal cell biomass, from 54.1 % to 140.57 %, with decreasing resolution (Fig. 6). Density RSD was more consistent and more closely aligned spatially with biomass than size and a had reduced mean cell RSD (27.59 % to 79.9 %, 5, column 2, (Supplementary Fig. 1).

3.3. Relationship between chlorophyll and copepod biomass at varying resolutions

At the smallest resolution (10 min bins, Fig. 3) there was a weak, significant relationship between chlorophyll and copepod biomass (Rs = 0.3, *p* < 0.001, n = 823). The relationship between chlorophyll and copepod biomass was tested at resolutions ranging from 0.05° to 0.9° increasing in steps of 0.01°. The strength of the relationship (Spearmans ρ) and the significance of the relationship are reported in Fig. 7. The relationship at the smallest spatial resolution (0.05° x 0.05°) was similar to the ten-minute bin (*p* < 0.001, n = 422). When decreasing resolution from 0.05° to 0.25°, there was little variation in the strength of the relationship and all relationships were significant. For lower resolutions, the strength and significance of the relationship between copepod biomass and chlorophyll became increasingly variable. For example, at a resolution of 0.83° the relationship was not significant and had weak positive correlation (ρ = 0.38, n = 11) while a resolution of 0.84° there was a strong positive correlation and the relationship was significant (ρ = 0.75, n = 11). For resolution lower than 0.9°, there were not enough data points (n < 10) to perform a Spearmans rank analysis (ideally n > 25).

3.4. Application to other variables.

The spatial distribution of chlorophyll concentrations is presented in Fig. 8 to demonstrate merging of other variables to a decreasing spatial resolution. Chlorophyll concentrations had a broader spatial pattern, where changes occurred over larger distances, than all copepod variables (Fig. 3A - C). These patterns are well captured in all selected resolutions (Fig. 8; Supplementary Fig. 2). There are no small-scale changes in chlorophyll concentration (Fig. 5) which was reflected in a lower, consistent RSD both spatially and across resolutions (Fig. 8; Supplementary Fig. 2). The area with the highest chlorophyll concentration, toward the southwest of the study area, also had the highest variation with adjacent cells, which was in turn reflected by a higher RSD (Fig. 8; Supplementary Fig. 2). Temperature and salinity (Supplementary
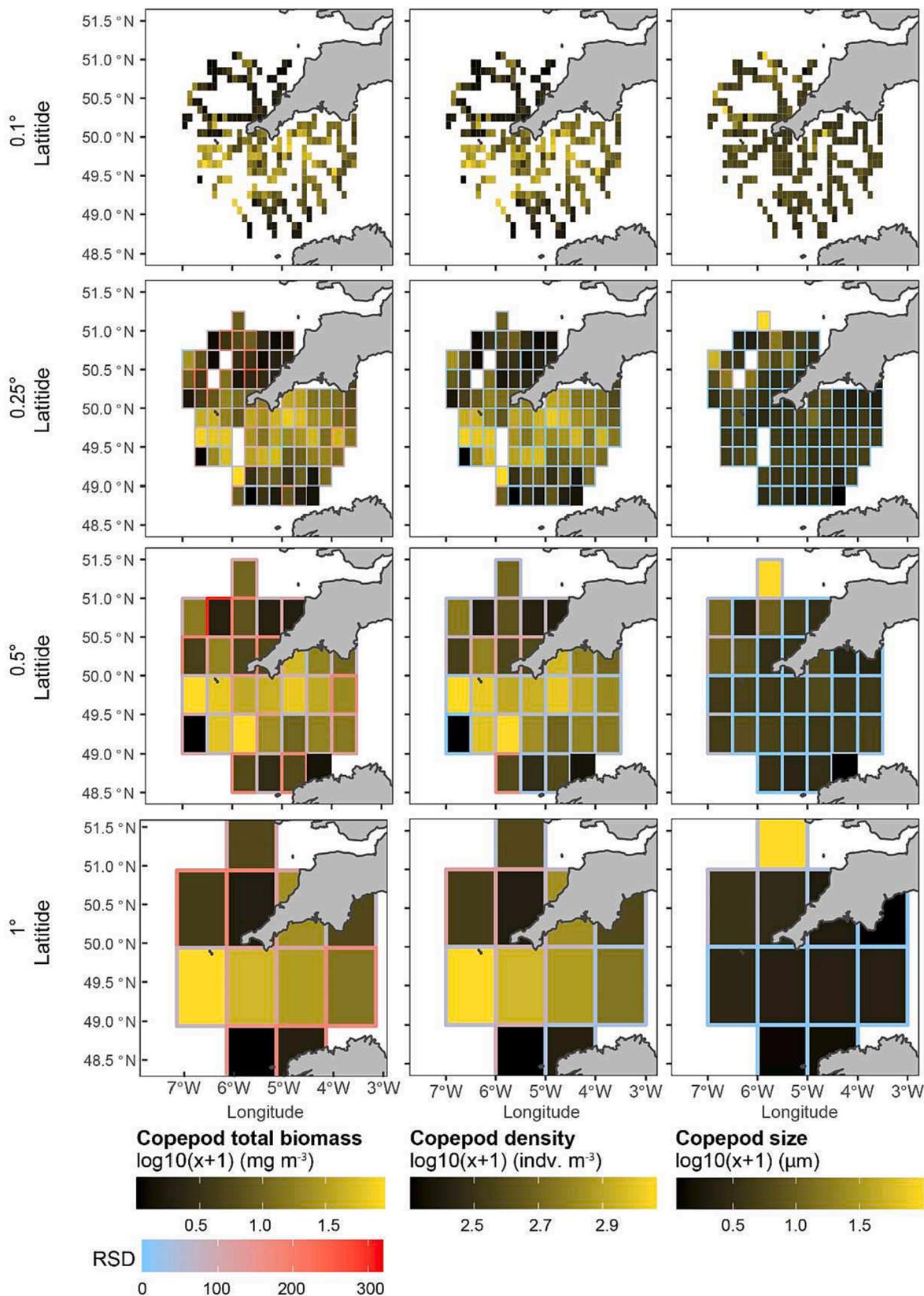
Fig. 5. 10 min bins merged to decreasing resolution for (column 1) copepod biomass (mg/m^{-3}($-|-$)) (log(x + 1)), (column 2) copepod density (indv. m^{-3}) (log(x + 1)) and (column 3) copepod size (μm)(log(x + 1)) for example resolutions (row 1) 0.1°, (row 2) 0.25°, (row 3) 0.5° and (row 4) 1°. Copepod color scales are the same as Fig. 3 for comparison. The cell border color indicates relative standard deviation (RSD, %) for the cell. Those cells without a border contain less than 3 data points. RSD color scale is the same for Figs. 5 and 8. For clarity RSD is also show alone for this figure in Supplementary Fig. 1.
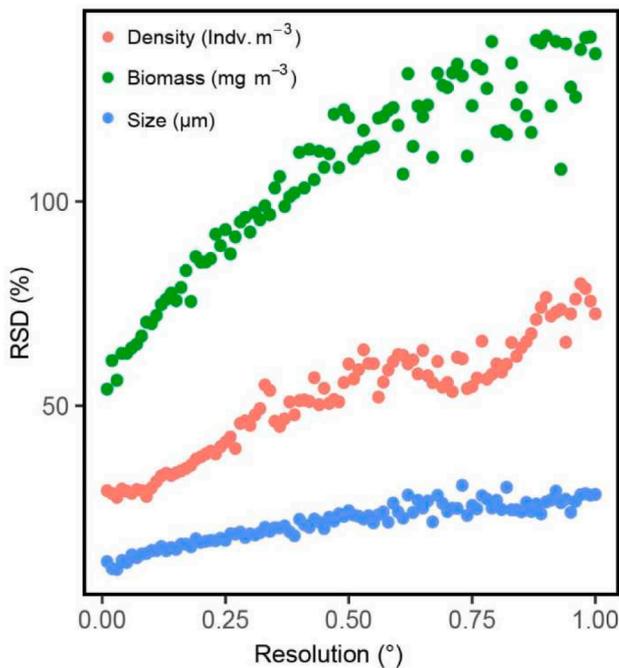
Fig. 6. Mean cell Relative Standard Deviation (RSD, %) for (green) copepod biomass (mg/m^{-3}(−|−)), (red) copepod density (indv. m^{-3}) and (blue) Size (µm), at all resolutions between 0.05° and 0.9° (increments = 0.01°). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
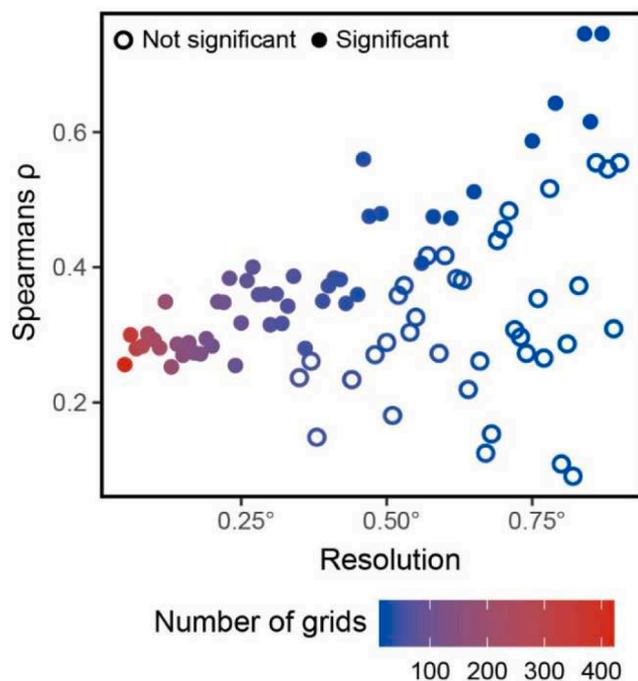

Fig. 7. The correlation between copepod biomass and chlorophyll using Spearmans ρ against decreasing spatial resolution. Resolutions decrease from 0.05° to 0.9° in increments of 0.01°. The number of grids (grid count) per resolution is indicated by the color scale. The significance of Spearmans correlation (*p* value < 0.05) is indicated by filled points where non-significant relationships are not filled, and significant relationships are filled.

materials) displayed similar results due to the absence of small spatial changes and broad, slower changes across the study area (Fig. 3E, 3F).

4. Discussion

The use of continuous instruments allowed for data to be obtained at small spatial scales, which in turn captured both wider spatial patterns and small-scale changes in copepod size, abundance and biomass. The small-scale changes in the copepod abundance, indicative of plankton patchiness (Mackas et al., 1985; Abraham, 1998), are not seen in the physical variables where patterns are study area wide. The physical oceanography of the Celtic Sea and Western Approaches, a seasonally stratified area, is well documented (Pingree et al., 1976; Pingree, 1980; Southward et al., 2004; Smyth et al., 2015). Stratification is known to influence plankton abundances (Fransz et al., 1984; Hure et al., 2022) but the absence of vertical data in this study does not allow for discussion of stratification or its influence on copepod abundance.

Neither surface temperature or salinity appeared correlated with copepod variables. However, the absence of a correlation between zooplankton and physical variables is in line with our understanding that small scale variations in the plankton are driven by a complex series of biological and physical interactions. These were reviewed by Atkinson et. al (2018), using a single point time series (L4 buoy) located in our study area. Average annual densities from the Continuous Plankton Recorder (Richardson et al., 2006) and reported by Johns (2006) find the majority of copepod families in lower abundance off the North coast of Cornwall. Although our data only cover 1 month, we find a similar spatial distribution, suggesting that the structure of zooplankton communities, within a specific area, remain similar both in time and space. The area wide patterns for copepod densities also match that of a previous study for the region using the PI (Pitois et al., 2021). Despite a lower taxonomic resolution obtained from image identification compared to microscope identification, another study using the PI found that the community structure described the PI is broadly in line with the L4 and CPR (Scott et al., 2021).

As machine learning classifiers for plankton identification from images collected with automated instruments improve in accuracy (The Turing Centre, 2021), it will be possible to discern zooplankton to increasing taxonomic resolution automatically. Added to the PI, this feature will allow for the removal of the subsampling step that is necessary when manually processing the images. Thus, it will be possible to obtain zooplankton data at higher taxonomic and spatial resolutions, quickly, and at a much lower cost compared to traditional methods. This will be a clear advantage of such systems. Although using the PI has the potential to yield an unprecedented spatial resolution, it cannot replicate the temporal resolution associated with devices such as CPR or longstanding time series such as L4. This is due to the PIs reliance on the vessel on which it's deployed, as it is unrealistic to expect a vessel to survey the same area repeatedly over long periods of time.

4.1. Optimizing survey design

Survey demands often result in ad-hoc, last minute changes reducing assurances of sampling the same spot at a consistent temporal resolution. Thus, a multi-method approach would yield the most complete description of the zooplankton. On the one hand, deploying plankton nets on vessels can help understand the vertical distribution of the plankton, whilst time series are invaluable to understand seasonal and long-term changes (Pitois and Yebra, 2022). On the other hand, understanding the small-scale fluctuations in the plankton, and what drives the high variation between neighboring water parcels, can be better understood using continuous data. Although here, data were subsampled and manually classified which limited the minimum achievable spatial resolution, the findings demonstrate the potential for these instruments to resolve these fine scale interactions driving variation. Furthermore, they demonstrate how the choice of resolution can affect
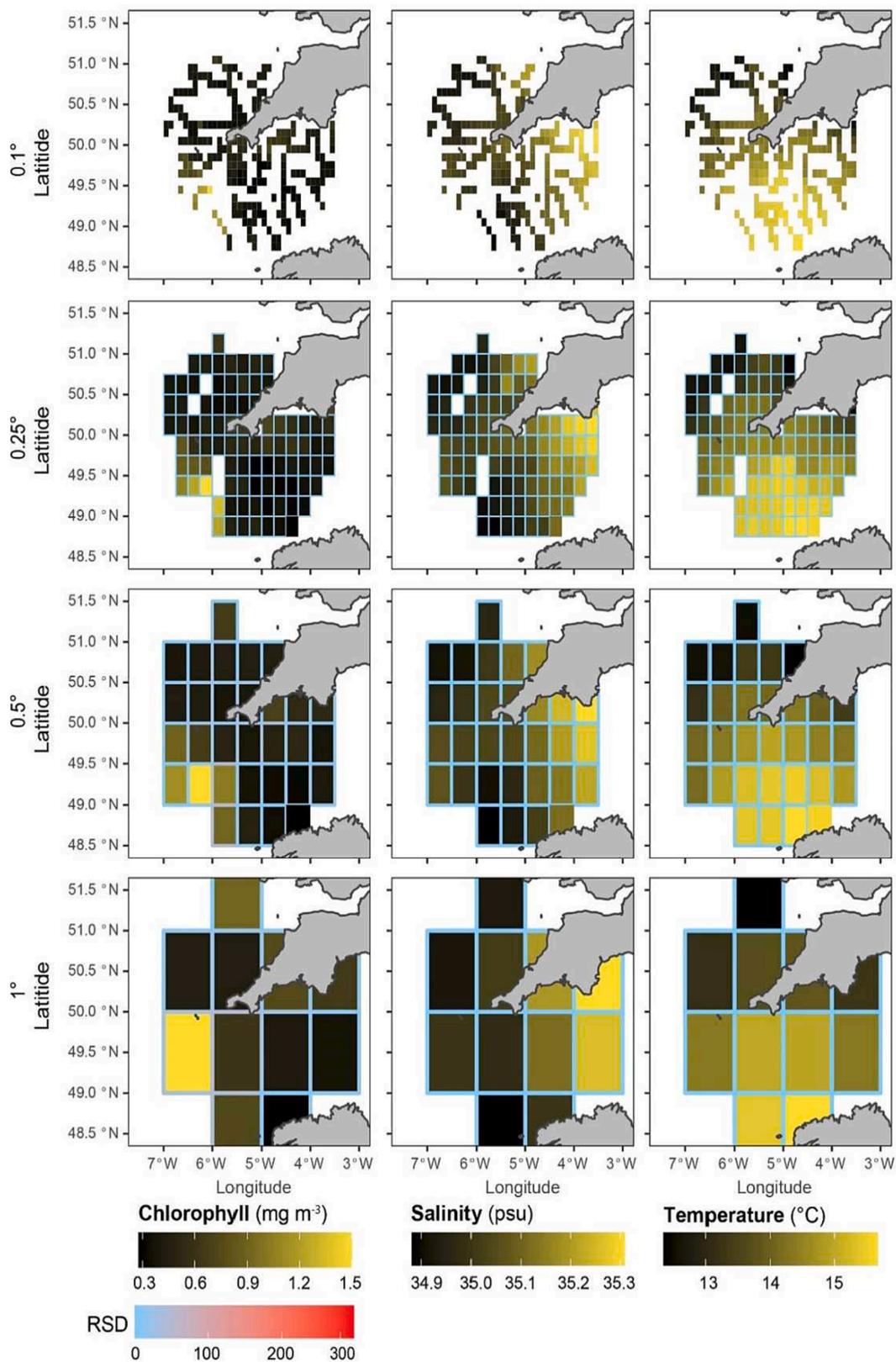
Fig. 8. 10 min bins merged to decreasing resolution for (column 1) chlorophyll (mg/m^{-3}($-$|-)), (column 2) sea surface temperature (°C) and (column 3) salinity (psu) for example resolutions (row 1) 0.1°, (row 2) 0.25°, (row 3) 0.5° and (row 4) 1°. Variable color scales are the same as Fig. 3 for comparison. The cell border color indicates relative standard deviation (RSD, %) for the cell. Those cells without a border contain less than 3 data points. RSD color scale is the same for Fig. 5. For clarity RSD is also shown alone for this figure in Supplementary Fig. 2.

the perceived picture of the plankton as well as relationships between plankton and related variables. Decreasing resolution can result in patterns being emphasized (e.g., chlorophyll) or small-scale changes being lost (e.g., copepod biomass). This demonstrates the 'risk' of a decreased sampling resolution in misrepresenting or incorrectly capturing trends. The variability within cells when merged to a decreasing resolution is not seen by an increased RSD, suggesting RSD is not sensitive to extreme values if the remainder of the merged cells are consistent. We can expect the changes in the data representation with decreased resolution to be reflected in statistical relationships between variables. Here, we chose to look at copepod biomass and chlorophyll concentrations. Chlorophyll data are readily available as a remote sensing package (Aumont et al., 2015) and many large-scale models rely on these data, inferring prey or carbon from chlorophyll (Landry, 1976; Carlotti and Poggiale, 2010). The relationships between chlorophyll and zooplankton are complex and reported relationships are inconsistent in the literature (Casini et al., 2008; Llope et al., 2012; Schultes et al., 2013; Giering et al., 2019). This variation partly stems from the different types of data and spatial temporal scales used between authors (Pitois et al., 2021). In our study, we find high variation in the strength and statistical significance of the relationship resulting only from changing spatial sampling resolution. Although all correlations are positive, we find both inconsistency in the significance and strength of the correlation at lower resolutions. It is likely that even finer resolution data, achieved through removal of subsampling, will yield the most accurate description of these relationships.

Sampling to the finest possible resolution may not be necessary or relevant to the survey's aim, but rather the choice of resolution, whether in space or time, should match the process studied. A sampling resolution too fine could incur unnecessary costs (in data storage and processing) and not be needed to accurately capture large-scale ecological patterns. For example, a coarser resolution than presented here (2° cells), has been used to successfully capture changes in copepod abundances over time as well investigate their relationship with various physical variables (Bedford et al., 2020). Conversely, too coarse a resolution may miss ecological processes that occur on scales finer than the selected sampling resolution. For example, collecting samples at a specific location once a year (temporal resolution) will not allow to capture seasonal variability.

For our descriptive study we find a spatial resolution of 0.25° to be a good compromise between capturing small scales changes and broader spatial patterns for copepod abundance and biomass. This resolution was the larger end of those resolutions that had consistency in the statistical relationship between copepod biomass and chlorophyll. Additionally, this resolution can be easily matched to existing products for modelling. For example, a remote sensing ecosystem model output has spatial resolution 0.25° x 0.25° (Aumont et al., 2015). Although each study will likely demand a different resolution dependent on the scale of the processes involved. Let's take, for example, a survey designed to study the timing and location of a specific fish spawning and the ecological processes affecting this (assuming both phyto- and zooplankton are variables measured). Prior knowledge will be used to select the overall location and timing of this survey as well as the parameters to measure. At this point, resolution of the measured processes and variables should be taken into consideration. If the survey occurs during the winter months when there is little activity in the plankton ecosystem, sampling these components at very fine temporal and / or spatial resolution is unlikely to be necessary. If, however, that survey occurs during the phytoplankton bloom, a time of fast change within the plankton ecosystem, then a finer resolution that matches the scale of these processes will need to be selected to accurately capture the changes. Similarly, sampling intensity can be adjusted during the course of the survey when and if changes are noticed.

Surveys tend to be designed to collect chosen parameters at pre-selected locations, usually as many as possible as can be covered by the survey based on time and budget available. In future, as automated tools become common place, optimizing survey design will need to combine different instruments that collect information complementarily to each other. For example, automated devices, such as the PI to collect surface data, alongside plankton nets to collect vertical data, would allow a more comprehensive description of the ecosystem studied. Automated tools, and the resolution they yield, may also help to quantify the variability associated with replicate tows resultant from plankton patchiness and be used to better understand its drivers (Wiebe et al., 1968; Lee and McAlice, 1979; Skjoldal et al., 2013). In theory, this could be achieved with the data presented here, by using a linear regression on the relationship between RSD and Resolution (Fig. 6). Although, it would be specific to this area and season, a 'survey snapshot' (Huret et al., 2018).

The small-scale resolution presented here, and the future potential for even finer resolution of biological parameters, achieved by reducing or eliminating subsampling, has only become recently possible with devices such as the PI. For this study, the time constraints associated with manually sorting images and the choice to subset the data to the study area meant only a small portion entire survey data were used (~2.4 %). This constraint does not apply for other variables, such as physical parameters, where no or little sample processing is required, and all data are instantly available. Recent improvement in classification algorithms (The Turing Centre, 2021) will bring the PI up to speed. Machine learning can eliminate the need for manual classification and thus the PI will be capable of providing continuous data at very fine resolutions (meters and minutes) where data do not need subsampling. Before these solutions can be implemented there are however several challenges that must be overcome, mainly related to the inability of the PI (and other similar systems) to process data as fast as they can collect it. These devices entail a phenomenal data collection rate. For example, if we were to use the full-size range the PI can image, we would collect up to 1 tb of data in<10 min. This currently not feasible as the technology or protocols to write images this fast does not yet exist. These devices clearly entail a phenomenal data collection rate. While the survey data totaled 2 tb for this study, we have made provision for 10 tb of data for the same survey to take place in 2023. The cost of storage and of compute is reducing but these must also form part of survey planning.

5. Conclusion

We demonstrate the importance of sampling resolution, in the context of pelagic studies, and how it affects relationships between selected measured parameters and their perceived resulting picture. The increasing use of automated and semi-automated technologies allows us to sample at a much finer resolution than previously possible across much larger spatial scales. This is especially true for zooplankton where the PI has the potential to provide unprecedented fine spatial data at moderate taxonomic resolution. Sampling resolution for each measured process will therefore need to be considered as part of optimum survey design. This is to ensure that sampling matches the resolution of the measured process at a specific place and time, and that only necessary data is collected to remain within the survey budgetary constraints. Integrating data collected from various instruments (both traditional and novel) will help to optimize sampling resolution for an improved understanding of ecosystem processes and ultimately, a more holistic view of marine ecosystems in all dimensions.

Author contributions

JS collected the data aboard the RV Cefas Endeavour. He led the writing of the manuscript and did all data analysis and created all figures. SP has led the work on the application of the PI system on-board RV Cefas Endeavour. She has managed the project, organized the field work and contributed to writing the manuscript. VC was responsible for the chlorophyll methodology, the conversion of fluorescence to chlorophyll and editing the manuscript. JT is the lead software developer for the PIA system and contributed to data management. PC is the lead scientist

responsible for the development of the PI technology and contributed to the writing of the manuscript. GM contributed to data interpretation and writing and editing the manuscript.

Declaration of Competing Interest

Authors Phil Culverhouse and Julian Tilbury are employed by Plankton Analytics Ltd which manufactor the Plankton Imager. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Data availability

Data will be made available on request.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.pocean.2022.102946.

References

Abraham, E.R., 1998. The generation of plankton patchiness by turbulent stirring. Nature 391, 577–580. https://doi.org/10.1038/35361.

Aumont, O., Ethé, C., Tagliabue, A., Bopp, L., Gehlen, M., 2015. PISCES-v2: An ocean biogeochemical model for carbon and ecosystem studies. Geosci. Model Dev. 8, 2465–2513. https://doi.org/10.5194/gmd-8-2465-2015.

Bean, T.P., Greenwood, N., Beckett, R., Biermann, L., Bignell, J.P., Brant, J.L., et al., 2017. a review of the tools used for marine monitoring in the UK: Combining historic and contemporary methods with modeling and socioeconomics to fulfill legislative needs and scientific ambitions. Front. Mar. Sci. 4, 263. https://doi.org/10.3389/fmars.2017.00263.

Beaugrand, G., Brander, K.M., Lindley, J.A., Souissi, S., Reid, P.C., 2003. Plankton effect on cod recruitment in the North Sea. Nature 426, 661–664. https://doi.org/10.1038/nature02164.

Bedford, J., Ostle, C., Johns, D.G., Atkinson, A., Best, M., Bresnan, E., et al., 2020. Lifeform indicators reveal large-scale shifts in plankton across the North-West European shelf. Glob. Chang. Biol. 26, 3482–3497. https://doi.org/10.1111/gcb.15066.

Carlotti, F., Poggiale, J.C., 2010. Towards methodological approaches to implement the zooplankton component in "end to end" food-web models. Prog. Oceanogr. 84, 20–38. https://doi.org/10.1016/j.pocean.2009.09.003.

Casini, M., Lövgren, J., Hjelm, J., Cardinale, M., Molinero, J. C., and Kornilovs, G., 2008. Multi-level trophic cascades in a heavily exploited open marine ecosystem. Proceedings of the Royal Society B: Biological Sciences 275, 1793–1801. http://doi.org/10.1098/rspb.2007.1752.

The Turing Centre (2021). DSG Report 2021.

Checkley, D.M., Dotson, R.C., Griffith, D.A., 2000. Continuous, underway sampling of eggs of Pacific sardine (Sardinops sagax) and northern anchovy (Engraulis mordax) in spring 1996 and 1997 off southern and central California. Deep Sea Res. 2 Top Stud. Oceanogr., 47, 1139–1155. http://doi.org/10.1016/S0967-0645(99)00139-3.

Commission, E., 2008. Directive 2003/6/EC of the European Parliament and the Council. Off. J. Eur. Union 22–142.

Culverhouse, P.F., Williams, R., Gallienne, C., Tilbury, J., Wall-Palmer, D., 2016. Ocean-Scale Monitoring of Mesozooplankton on Atlantic Meridional Transect 21. J. Mar. Biol. Aquaculture 2, 1–13. https://doi.org/10.15436/2381-0750.16.018.

Danovaro, R., Carugati, L., Berzano, M., Cahill, A.E., Carvalho, S., Chenuil, A., et al., 2016. Implementing and innovating marine monitoring approaches for assessing marine environmental status. Front. Mar. Sci. 3, 213. https://doi.org/10.3389/fmars.2016.00213.

de Moustier, C., 1986. Beyond bathymetry: Mapping acoustic backscattering from the deep seafloor with sea beam. J. Acoust. Soc. Am. 79, 316–331. https://doi.org/10.1121/1.393570.

Doray, M., Petitgas, P., Romagnan, J.B., Huret, M., Duhamel, E., Dupuy, C., et al., 2018. The PELGAS survey: Ship-based integrated monitoring of the Bay of Biscay pelagic ecosystem. Prog. Oceanogr. 166, 15–29. https://doi.org/10.1016/j.pocean.2017.09.015.

Fransz, H.G., Miquel, J.C., Gonzalez, S.R., 1984. Mesozooplankton composition, biomass and vertical distribution, and copepod production in the stratified central north sea. Neth. J. Sea Res. 18, 82–96. https://doi.org/10.1016/0077-7579(84)90026-7.

Giering, S.L.C., Wells, S.R., Mayers, K.M.J., Schuster, H., Cornwell, L., Fileman, E.S., et al., 2019. Seasonal variation of zooplankton community structure and trophic position in the Celtic Sea: A stable isotope and biovolume spectrum approach. Prog. Oceanogr. 177, 101943 https://doi.org/10.1016/j.pocean.2018.03.012.

Heath, M.R., 2005. Regional variability in the trophic requirements of shelf sea fisheries in the Northeast Atlantic, 1973–2000. ICES J. Mar. Sci. 62, 1233–1244. https://doi.org/10.1016/j.icesjms.2005.04.010.

Hure, M., Batistić, M., Garić, R., 2022. Copepod Diel Vertical Distribution in the Open Southern Adriatic Sea (NE Mediterranean) under Two Different Environmental Conditions. Water 14. https://doi.org/10.3390/w14121901.

Huret, M., Bourriau, P., Doray, M., Gohin, F., Petitgas, P., 2018. Survey timing vs. ecosystem scheduling: Degree-days to underpin observed interannual variability in marine ecosystems. Prog. Oceanogr. 166, 30–40. https://doi.org/10.1016/j.pocean.2017.07.007.

ICES, 2015. Manual for International Pelagic Surveys (IPS), in: Surveys, W.G.o.I.P. Copenhagen.

Johnson, M., De Soto, N.A., Madsen, P.T., 2009. Studying the behaviour and sensory ecology of marine mammals using acoustic recording tags: A review. Mar. Ecol. Prog. Ser. 395, 55–73. https://doi.org/10.3354/meps08255.

Johnson, M.P., Tyack, P.L., 2003. A digital acoustic recording tag for measuring the response of wild marine mammals to sound. IEEE J. Ocean. Eng. 28, 3–12. https://doi.org/10.1109/JOE.2002.808212.

Kupschus, S., Schratzberger, M., Righton, D., 2016. Practical implementation of ecosystem monitoring for the ecosystem approach to management. J. Appl. Ecol. 53, 1236–1247. https://doi.org/10.1111/1365-2664.12648.

Lampert, W., 1989. The Adaptive Significance of Diel Vertical Migration of Zooplankton. Funct. Ecol. 3, 21. https://doi.org/10.2307/2389671.

Landry, M.R., 1976. The structure of Marine Ecosystems: An alternative. Harvard University Press. https://doi.org/10.1007/BF00386669.

Lauria, V., Attrill, M.J., Brown, A., Edwards, M., Votier, S.C., 2013. Regional variation in the impact of climate change: Evidence that bottom-up regulation from plankton to seabirds is weak in parts of the Northeast Atlantic. Mar. Ecol. Prog. Ser. 488, 11–22. https://doi.org/10.3354/meps10401.

Lee, W.Y., McAlice, B.J., 1979. Sampling variability of marine zooplankton in a tidal estuary. Estuar. Coast. Mar. Sci. 8, 565–582. https://doi.org/10.1016/0302-3524(79)90036-7.

Llope, M., Licandro, P., Chan, K.S., Stenseth, N.C., 2012. Spatial variability of the plankton trophic interaction in the North Sea: A new feature after the early 1970s. Glob. Chang. Biol. 18, 106–117. https://doi.org/10.1111/j.1365-2486.2011.02492.x.

Lombard, F., Boss, E., Waite, A.M., Uitz, J., Stemmann, L., Sosik, H.M., et al., 2019. Globally consistent quantitative observations of planktonic ecosystems. Front. Mar. Sci. 6 https://doi.org/10.3389/fmars.2019.00196.

Mackas, D.L., Denman, K.L., Abbott, M.R., 1985. Plankton patchiness: biology in the physical vernacular. Bull. Mar. Sci. 37, 653–674.

Mann, D.A., Hawkins, A.D., Jech, J.M., 2008. Active and Passive Acoustics to Locate and Study Fish. Fish Bioacoustics (Springer) 279–309. https://doi.org/10.1007/978-0-387-73029-5_9.

Olson, R.J., Zettler, E.R., DuRand, M.D., 2018. Phytoplankton analysis using flow cytometry. In: Handbook of methods in aquatic microbial ecology. CRC Press, pp. 175–186.

Owens, N.J.P., 2014. Sustained UK marine observations. Where have we been? Where are we now? Where are we going? Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. 372, 20130332. https://doi.org/10.1098/rsta.2013.0332.

Petersen, W., Colijn, F., 2017. FerryBox Whitebook.

Pingree, R.D., 1980. Physical Oceanography of the Celtic Sea and English Channel. Elsevier Oceanogr. Ser. 24, 415–465. https://doi.org/10.1016/S0422-9894(08)71358-8.

Pingree, R.D., Holligan, P.M., Mardell, G.T., Head, R.N., 1976. The influence of physical stability on spring, summer and autumn phytoplankton blooms in the celtic sea. J. Mar. Biol. Assoc. U. K. 56, 845–873. https://doi.org/10.1017/S0025315400020919.

Pitois, S.G., Tilbury, J., Bouch, P., Close, H., Barnett, S., Culverhouse, P.F., 2018. Comparison sof a Cost-Effective Integrated Plankton Sampling and Imaging Instrument with Traditional Systems for Mesozooplankton Sampling in the Celtic Sea. Front. Mar. Sci. 5, 1–15. https://doi.org/10.3389/fmars.2018.00005.

Pitois, S.G., Graves, C.A., Close, H., Lynam, C., Scott, J., Tilbury, J., et al., 2021. A first approach to build and test the Copepod Mean Size and Total Abundance (CMSTA)

ecological indicator using in-situ size measurements from the Plankton Imager (PI). Ecol. Ind. 123, 107307 https://doi.org/10.1016/j.ecolind.2020.107307.

Pitois, S., Yebra, L., 2022. Contribution of marine zooplankton time series to the United Nations Decade of Ocean Science for Sustainable Development. ICES J. Mar. Sci. 79, 722–726.

Postel, L., Fock, H., Hagen, W., 2000. Biomass and abundance. In: ICES Zooplankton Methodology Manual. Elsevier, pp. 83–192. https://doi.org/10.1016/b978-012327645-2/50005-0.

Richardson, A.J., Walne, A.W., John, A.W.G.G., Jonas, T.D., Lindley, J.A., Sims, D.W., et al., 2006. Using continuous plankton recorder data. Prog. Oceanogr. 68, 27–74. https://doi.org/10.1016/j.pocean.2005.09.011.

Schultes, S., Sourisseau, M., Le Masson, E., Lunven, M., Marié, L., 2013. Influence of physical forcing on mesozooplankton communities at the Ushant tidal front. J. Mar. Syst. 109–110, S191–S202. https://doi.org/10.1016/j.jmarsys.2011.11.025.

Scott, J., Pitois, S., Close, H., Almeida, N., Culverhouse, P., Tilbury, J., et al., 2021. In situ automated imaging, using the Plankton Imager, captures temporal variations in mesozooplankton using the Celtic Sea as a case study. J. Plankton Res. 43, 300–313. https://doi.org/10.1093/plankt/fbab018.

Simmonds, J., MacLennan, D.N., 2008. Fisheries acoustics: theory and practice. John Wiley & Sons.

Skjoldal, H.R., Wiebe, P.H., Postel, L., Knutsen, T., Kaartvedt, S., Sameoto, D.D., 2013. Intercomparison of zooplankton (net) sampling systems: Results from the ICES/GLOBEC sea-going workshop. Prog. Oceanogr. 108, 1–42. https://doi.org/10.1016/j.pocean.2012.10.006.

Smyth, T., Atkinson, A., Widdicombe, S., Frost, M., Allen, I., Fishwick, J., et al., 2015. The Western Channel Observatory. Prog. Oceanogr. 137, 335–341. https://doi.org/10.1016/j.pocean.2015.05.020.

Southward, A.J., Langmead, O., Hardman-Mountford, N.J., Aiken, J., Boalch, G.T., Dando, P.R., et al., 2004. Long-term oceanographic and ecological research in the western English Channel. http://doi.org/10.1016/S0065-2881(04)47001-1.

Steinberg, D.K., Goldthwait, S.A., Hansell, D.A., 2002. Zooplankton vertical migration and the active transport of dissolved organic and inorganic nitrogen in the Sargasso Sea. Deep Sea Res. 1 Oceanogr. Res. Pap. 49, 1445–1461. https://doi.org/10.1016/S0967-0637(02)00037-7.

Steinberg, D.K., Landry, M.R., 2017. Zooplankton and the Ocean Carbon Cycle. Ann. Rev. Mar. Sci. 9, 413–444. https://doi.org/10.1146/annurev-marine-010814-015924.

Strickland, J. D., Parsons, T.R., 1972. 1972. A practical handbook of seawater analysis. 2nd ed. Bull. Fish. Res. Bd. Can, 167.

Taylor, A.H., Allen, J.I., Clark, P.A., 2002. Extraction of a weak climatic signal by an ecosystem. Nature 416, 629–632. https://doi.org/10.1038/416629a.

Wiebe, P.H., Benfield, M.C., 2003. From the Hensen net toward four-dimensional biological oceanography. Prog. Oceanogr. 56, 7–136. https://doi.org/10.1016/S0079-6611(02)00140-4.

Wiebe, P.H., Wiebe, P.H., Holland, W.R., 1968. Plankton Patchiness: Effects on Repeated Net Tows. Limnol. Oceanogr. 13, 315–321. https://doi.org/10.4319/lo.1968.13.2.0315.