



J. Plankton Res. (2021) 43(3): 475–491. First published online May 18, 2021 doi:10.1093/plankt/fbab037

ORIGINAL ARTICLE

Use of optical imaging datasets to assess biogeochemical contributions of the mesozooplankton

AMY E. MAAS*, HANNAH GOSSNER, MAISIE J. SMITH AND LEOCADIO BLANCO-BERCIAL

BERMUDA INSTITUTE OF OCEAN SCIENCES, 17 BIOLOGICAL STATION, ST. GEORGES, GEO1 ST. GEORGE'S, BERMUDA

*CORRESPONDING AUTHOR: amy.maas@bios.edu

Received September 11, 2020; editorial decision April 26, 2021; accepted April 26, 2021

Corresponding editor: Xabier Irigoien

The increasing use of image-based observing systems in marine ecosystems allows for more quantitative analysis of the ecological zonation of zooplankton. Developing methods that take advantage of these systems can provide an increasingly nuanced understanding of how morphometric characteristics (especially size) are related to distribution, abundance and ecosystem function via a wider application of allometric calculations of biogeochemical fluxes. Using MOCNESS sampling of zooplankton near the Bermuda Atlantic Time Series and a ZooSCAN/EcoTaxa pipeline, we apply a new taxonomically resolved biomass to biovolume dataset and a suite of R scripts that provide information about the relationships between zooplankton size, taxonomy, distribution, depth of migration, magnitude of migration and biogeochemical contributions (e.g. respiratory O₂ consumption). The analysis pipeline provides a framework for quantitatively comparing and testing hypotheses about the distribution, migration patterns and biogeochemical impacts of mesozooplankton. Specifically, our code helps to visualize a size-based structure in the extent of vertical migration and allow for a quantification of the relative importance of non-migratory versus migratory organisms of various size classes. It additionally allows us to quantify the error associated with various methods of calculating active flux, with size-based analysis being the most important methodological choice, and taxonomic identification being the least.

KEYWORDS: allometry; ZooSCAN; MOCNESS; diel vertical migration

INTRODUCTION

The structure of the planktonic food web plays a fundamental role in regulating the export of organic carbon, nitrogen and phosphorus to the deep ocean (the Biological Pump; Ducklow *et al.*, 2001, Longhurst & Harrison, 1989). Migratory species consume food in the warm, oxygenated, photic zone during the night and move to deeper, cooler layers by day. This diel vertical migration (DVM) transfers atmospheric CO₂, fixed by phytoplankton into organic carbon and consumed by zooplankton in surface waters, to ocean depths where there is potential for long term sequestration of the organic carbon while also fueling the mesopelagic microbial community energy demand (Longhurst *et al.*, 1990, Maas *et al.*, 2020, Steinberg *et al.*, 2008b). Similarly, through the excretion of nitrogenous compounds at depth, DVM diminishes the availability of this limiting nutrient for phytoplankton in near-surface waters, greatly shaping the capacity for new production (Longhurst & Harrison, 1988). Globally, DVM accounts for 15–40% of the total organic carbon export from the surface to ocean interior (Bianchi *et al.*, 2013, Steinberg *et al.*, 2000), exceeding in many cases the carbon flux associated with passively sinking particles (Kobari *et al.*, 2013, Kobari *et al.*, 2008, Steinberg *et al.*, 2008b). Similar percentages have been reported for the nitrogen-related processes (Al-Mutairi & Landry, 2001, Dam *et al.*, 1995, Schnetzer & Steinberg, 2002).

A review of the published literature on zooplankton contributions to the biological pump by Steinberg and Landry (2017) shows that there is still significant uncertainty in our understanding of the drivers of this process, with the basic relationship between biomass and the movement of carbon by these vertically migratory zooplankton (active flux) still unresolved. We posit that biomass alone does not adequately describe the zooplankton community because it is a product of both the size and abundance of the organisms. These two factors respond differently to environmental variables; for instance, across large geographical provinces conditions arise where environmental forcing generates communities of fewer, larger individuals having the same biomass as communities with many but smaller individuals. Since physiological contributions to biogeochemistry are size dependent (smaller organisms have a larger respiratory contribution per unit body mass; Ikeda, 2014b) and most trophic dynamics (such as prey choice and grazing rate; Hansen *et al.*, 1997, Straile, 1997) are generally related to organismal size, this results in communities of the same biomass having distinct biogeochemical signatures.

Beyond the physiological implications of size class distribution of migratory organisms, the size spectra of the community also change the vertical extent of

migration. Recent analyses of CalCOFI data quantitatively demonstrate that intermediate-sized copepod species have stronger DVM behavior (Ohman & Romagnan, 2016), while the smallest and largest copepod species refrain from migrating. Within those intermediate classes, larger species show a greater amplitude in their DVM. Detailing the relationships between DVM extent, size class and midwater hydrographic parameters is important because models have suggested that the depth and intensity of DVM have substantial implications for the efficiency of active flux (Archibald *et al.*, 2019), and consequently, the biological pump. The depth of the migration is important not only because it removes the possibility of the respired CO₂ to be returned to the atmosphere by mixing, but also because deeper waters tend to be colder. Since most plankton are endotherms, their physiology is directly influenced by the temperature of their environment, making their metabolism slow as they move deeper into the water column—reducing their active flux. Deeper water can, moreover, have reduced oxygen concentrations that influence the physiology of migratory organisms to further reduce respiratory flux (Kiko & Hauss, 2019, Seibel, 2011).

Additionally, efficiency of total export is modified by the environment and community composition in the midwater (~200–1000 m; the region below the euphotic zone). In this ‘twilight zone’ migratory zooplankton and stationary resident species assemblages repackage and transform surface derived flux in a process associated with flux attenuation (Robinson *et al.*, 2010, Steinberg & Landry, 2017, Stukel *et al.*, 2019). The modification of flux to depth depends on zooplankton community structure, biomass, and the behavioral and physiological consequences of the midwater environment; however, there is a lack of detailed understanding of these midwater communities. This is partially due to undersampling, but it is also an artifact of historical sampling methods, which made the generation of datasets that considered the taxonomy and size of individual zooplankton intensely laborious.

As imaging systems and informatics pipelines improve; however, there are greater opportunities to use size-based assessments of the zooplankton community to more accurately describe the structure of migratory and midwater assemblages (Lombard *et al.*, 2019, Noyon *et al.*, 2020, Ohman *et al.*, 2019). Based on size class information, allometric equations can then be applied to describe how both resident and migratory zooplankton are likely to contribute to biogeochemical cycling (Kiko *et al.*, 2020, Kiko & Hauss, 2019). The objective of this paper is to demonstrate a set of R codes that can be used to compare various methods for calculating the biogeochemical contributions of zooplankton to active flux, to explore

Table I: MOCNESS net tow metadata. The top net (0–50) typically captured the mixed layer, while the second net (50 ~200 m) captured the deep chlorophyll a max

Cruise	Year	Month	Day	D/N	Cast #	Lat	Long	Net depths (m)
AE1614	2016	July	11	N	3	32°10.930 N	64°28.521 W	0, 50, 225, 300, 450, 600, 750, 900, 1 000
			11	D	4	32°10.846 N	64°28.811 W	0, 50, 225, 300, 450, 600, 750, 900, 1 000
AE1712*	2017	July	7	D	8	31°42.347 N	64°12.654 W	0, 50, 200, 275, 400, 550, 700, 850, 1 000
			11	N	9	31°42.161 N	64°12.941 W	0, 50, 200, 275, 400, 550, 700, 850, 1 000
AE1819	2018	July	3	D	10	31°39.961 N	64°09.535 W	0, 50, 175, 250, 400, 550, 700, 850, 1 000
			4	N	11	31°40.133 N	64°10.212 W	x, 50, 175, 250, 400, 550, 700, 850, 1 000
AE1830	2018	October	28	N	12	31°49.798 N	64°03.727 W	0, 50, 200, 300, 400, 550, 700, 900, 1 000
			30	D	13	31°48.319 N	64°07.294 W	0, 50, 200, 300, 400, 550, 700, 900, 1 000

The second net from the bottom typically captured the local oxygen minima. The only tow with validated images is AE1712 (*). The top net from the night time in July 2018 failed (x), so this net or day/night pair was excluded from some analyses.

particle size distribution patterns in zooplankton, and to quantitatively relate the depth of migration to the size of the migrator using ZooSCAN and allometric approaches. We use four paired day/night tows from the Bermuda Atlantic Time Series (BATS) site to demonstrate this code.

METHODS

MOCNESS net sampling

To obtain samples, a 1-m² Multiple Opening/Closing Net and Environmental Sensing System (MOCNESS; Wiebe *et al.*, 1985) equipped with 150- μ m nets was deployed during the mid-day and mid-night on cruises carried out in the vicinity of the Bermuda Atlantic Time-series Study site (BATS) in July of 2016, 2017 and 2018 as well as October 2018 (eight casts in total, 63 discrete nets; Table I). During the July 2018 cruise the top net of the night cast failed, and consequently, some analyses exclude this pair of tows. The net sampling plan used adaptive profiling, modifying the depth of closure to capture eight distinct ecological zones from surface to 1000 m depth: e.g. the thermocline, deep chl-*a* maximum (DCM), above the O₂ minimum, within the O₂ minimum core and below the O₂ minimum (Maas *et al.*, 2014, Steinberg *et al.*, 2008a). The actual tow depths were set on a cruise-by-cruise basis, using CTD profiles to determine the depths of each ecological zone. Upon retrieval, the catch from each of the eight discrete nets were divided into splits. Half were preserved in 95% undenatured ethanol and the remainder preserved in buffered 4% formalin in seawater.

Image analysis

To characterize the abundance and size class distributions of organisms, a representative subsample of the zooplankton community from each net were imaged from the formalin-preserved samples, then measured and broadly taxonomically classified, using a ZooSCAN ver. 3 at 4800 dpi (following the methods in: Gorsky *et al.*, 2010, Vandromme *et al.*, 2012). In order to better represent all size classes in the images, the original sample was divided into three size categories. All individuals larger than 2 cm were selected by eye and scanned separately from all the others. The remainder of the sample was sieved through a 1-mm mesh sieve, and both size fractions were individually scanned. From these smaller two size fractions, at least 1500 particles were scanned after subsampling using a Motoda splitter (Motoda, 1959), requiring generation of two separate scans for both the medium and small size classes. This resulted in a total of five images per net. Raw images were then processed in ZooProcess (Gorsky *et al.*, 2010, Vandromme *et al.*, 2012), and the resulting vignettes, along with a measurement and metadata file, were uploaded to EcoTaxa (<https://ecotaxa.obs-vlfr.fr/>; Picheral *et al.*, 2017) for machine-assisted identification, using a training set developed by the authors. One paired set of tows was completely validated (July 2017) for taxonomy. For the remaining tows image validation was not completed, and downstream analyses relied on Ecotaxa prediction of live versus dead. Assessment of the predictive power of the 2017 training dataset (predicted vs. validated) for each taxonomic grouping is currently not possible in the program, but we were able to determine that our training set gives efficacy

values of 83% correct ID of living and 93% correct ID of dead particles.

Biomass and biovolume analysis

We conducted an analysis of the estimated biovolume of a particle to its measured mass for a suite of taxa to allow for the application of both taxon-specific and generalized equations to our datasets. This provides a direct relationship between the optically estimated measurements of size calculated by the ZooProcess/EcoTaxa pipeline and biomass (wet mass, WM; dry mass, DM). 918 individual zooplankton encompassing a range of taxa and sizes were scanned following the previously described imaging pipeline. Each individual organism was then weighed wet on a Mettler-Toledo XPR microbalance, dried for 3 days in a drying oven at 60°C, and reweighed. Image-based biovolume (mm³) measurements were related to dry weight (mg) measurements by setting the intercept to 0 and assuming an ellipsoidal shape where:

$$mm^3 = \frac{4}{3} \times \pi \times (minor \times 0.5)^2 \times (major \times 0.5) \quad (1)$$

Regressions and statistical significance were tested using SPSS version 27. We used biovolume for our subsequent conversion factor, although area, ESD and ferret are alternate metrics from the literature (Forest *et al.*, 2012, Lehet & Hernández-León, 2009, Ohman & Romagnan, 2016). The full dataset is provided (Supplementary Data 1) to allow cross-comparison.

Setup of the R code

The R code was designed to read in a spreadsheet directly downloaded from EcoTaxa, as well as hydrographic data from the MOCNESS sampling system (specifically temperature and depth values), then put the data into a format that allows for application of any set of biomass or allometric equations as well as specific visualization (Figure 1). The code is modular and annotated so that researchers can tune the equations to their specific research interests or taxonomic focus. Datasets that are the direct product of a ZooSCAN (i.e. without the classification step) can also be used in case they were done exclusively on hand-picked individuals of interest, but we strongly suggest that an initial classification of particles to live/dead be conducted, as it is our experience that ~50% of the images in whole-community scans are threads, detritus, bubbles, shadows, etc.

The first step of the R code (Reformatting; Figure 1) reads the.tsv file containing the data of each particle and selects relevant metadata, measurements and

taxonomy. For our data we use the ‘object_area’, ‘object_major’, and ‘object_minor’. Additional metrics such as ‘object_ESD’ can be included if it suits the data analysis. To make quantitative m³ and m² calculations the ‘object_depth_min’, ‘object_depth_max’, ‘sample_tot_vol’ and ‘acq_sub_part’ (split) are required. The code then converts the ‘object_ID’ into the relevant information (typically cruise, tow number, net and size fraction). All of the tows and nets are then assigned a header for charts and graphs and additional relevant environmental metadata. For our analysis, we included day/night and median net temperature, although other information can also be assigned based on the scientific question (e.g. median oxygen of a net, day length, etc.).

The EcoTaxa measurements are in pixels, so a set of calculations must be performed to give length and biovolume in mm. For our analysis, individual biovolumes were calculated assuming an ellipsoidal shape as described previously, and then the taxon-specific biovolume to DM conversion is applied. If the taxonomic assignment does not fall into one of the pre-defined categories, it assumes the conversion factor of a calanoid copepod. To assist in particle size distribution (PSD) binning, biovolumes are log transformed during this step. Additional factors, including taxonomy, as well as the number of splits, volume filtered, net temperature and day/night designation are assigned to each particle during this step. At the end of this formatting step, there is a.csv file for each MOCNESS with all the metadata, taxonomy and measurements in a standard format ready for downstream analysis.

The next step of the code (Binning; Figure 1) makes a summary of the particles. In this step, the bin size is chosen by the user. For this example, we set our bins to span from 0.001 and 1000 mm³ because of our observed particle sizes, and separated by 0.25 on the log scale (this means each bin is 1.778 bigger than the previous). During this step, user-selected allometric equations used to assess biogeochemical processes are applied. The code also allows for taxonomically specific equations if desired. For this demonstration, we chose to calculate respiratory active flux using a consistent set of allometric scaling, temperature (Q_{10}) and respiratory quotient (RQ) relationships. Specifically, oxygen utilization (μmol O₂ ind⁻¹ h⁻¹) of each ‘live’ particle was calculated based on its biomass (DM; mg) and the average temperature (°C) of the net from which it was collected using the following equation:

$$\text{Oxygen Consumption Rate} = \text{EXP}(a_0 + (a_1 * \ln(\text{DM})) + (a_2 * (\text{TempC}))) / 22.4 \quad (2)$$

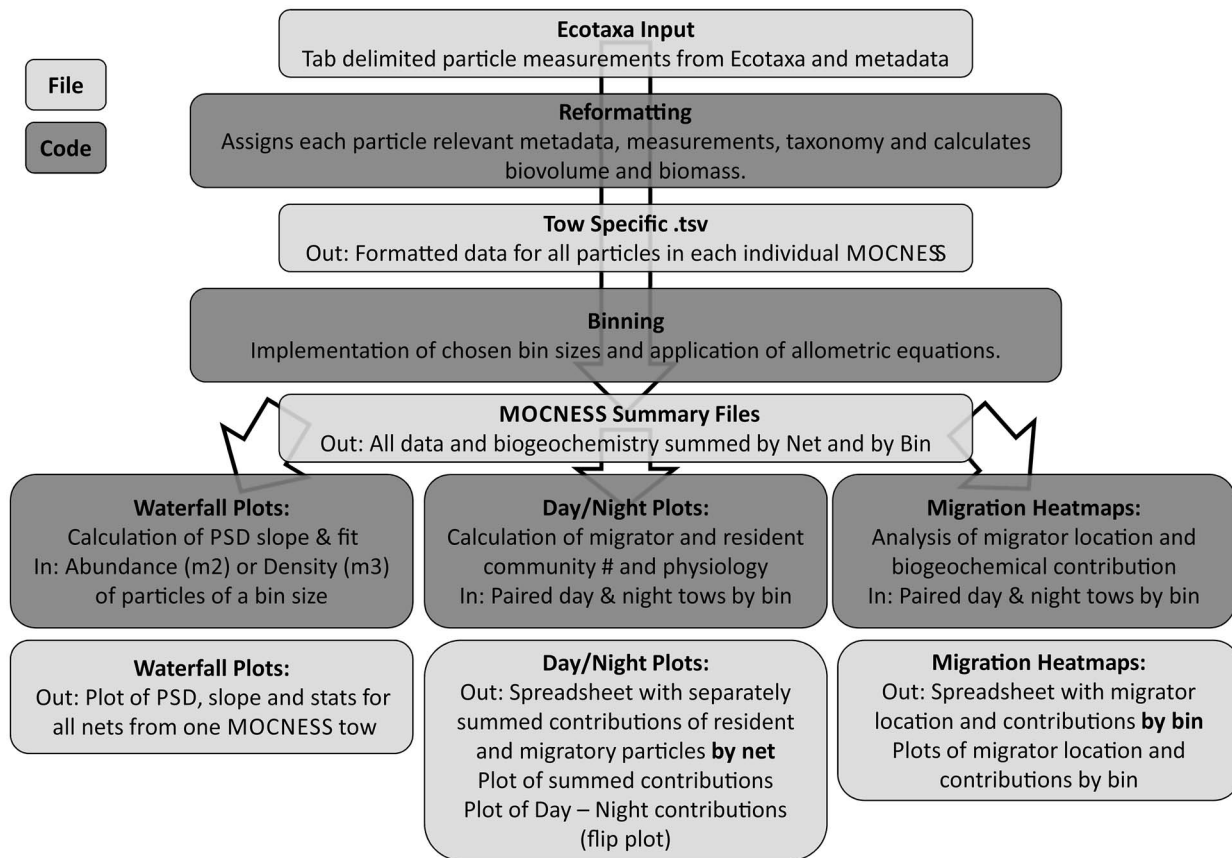


Fig. 1. Workflow and outcomes of the R code.

We used taxon-specific coefficients from Ikeda (Ikeda, 2013a, 2013b, 2014a, Ikeda *et al.*, 2001, Ikeda & Takahashi, 2012), since his global zooplankton model (Ikeda, 2014b) assumes no taxon-specific effects of temperature or body size (Table II). Oxygen consumption rates were transformed to respiratory CO_2 using a respiratory quotient (RQ) and applied exclusively to particles that are defined as representing migratory individuals, to generate an estimate of the active transport. We applied the taxon-specific RQ from Mayzaud *et al.* (2005; Table II). If the taxonomic assignment did not fall into one of the pre-defined categories the code assumes the coefficients of a copepod. This decision was based on the recent finding that copepod-specific allometric scaling factors were the best predictor of measured metabolic rates during experiments run the Northwest Pacific {Maas, 2021 #2264}. Other equations, including nitrogen excretion (Ikeda, 2014b), mortality (Hirst & Kiørboe, 2002, Zhang & Dam, 1997), DOC (Maas *et al.*, in review, Steinberg *et al.*, 2000), or any other physiological process could be similarly applied. Additional environmental

modifications such as day length, depth of capture or oxygen availability could be incorporated into the biogeochemical calculations, if regionally relevant, using implementation of an additional lookup table as was applied here for temperature.

From these data, we create three different products. The first is an analysis of the particle size class distribution (PSD), visualized with waterfall plots. The second and third are characterizations of the abundance and physiological contributions of the migratory and resident communities either summed by net or based on bin size. These are visualized as vertically stratified bar plots, a flip plot showing the location of migratory organisms and a heatmap showing the daytime and nighttime location of migratory organisms. These calculations are made by assuming that when daytime abundances are subtracted from night time these provide an estimate of the migratory community. The smaller of the paired day and night abundances provides a rough estimate of the resident midwater species within a size class at a particular depth interval, although these

Table II: Taxon-specific respiratory quotient (RQ), allometric and temperature scaling coefficients. RQ from Mayzaud *et al.* (2005)

Taxonomic Group	RQ	a ₀	a ₁	a ₂	Source
Amphipod	1.35	0.407	0.743	0.037	Ikeda 2013a
Chaetognath	1.35	−0.173	0.805	0.068	Ikeda & Takahashi 2012
Copepod	0.87	−0.399	0.801	0.069	Ikeda <i>et al.</i> 2001
Euphausiid	1.35	0.392	0.753	0.046	Ikeda 2013b
Mollusca	0.94	−0.56	0.82	0.046	Ikeda 2014a

Coefficients are from taxon-specific models (Ikeda, 2013a, Ikeda, 2013b, Ikeda, 2014a, Ikeda *et al.*, 2001, Ikeda & Takahashi, 2012). If the taxonomy of the particle was not explicitly listed the code applies copepod-specific values.

measurements do not account for reverse migration or general patchiness.

The R code associated with the analysis pipeline can be found in [Supplementary file 2](#) and both it and the particle dataset are available on Github at <https://github.com/blancoercial/BiogeochemicalContributions>. The output figures for all nets from the full code are provided in [Supplementary file 3](#).

Active flux calculations (post-processing the output)

To calculate active flux from the output of the R code, data from the.csv need to be summed based on a depth for flux chosen by the user. Typically, respiratory active transport has been calculated as the difference between the allometrically derived physiological contributions of the nighttime and daytime biomass in the euphotic zone (EZ), corrected to the midwater temperature of the average depth of migration. This can be implemented in the code by setting the temperature of all of the EZ nets to the average midwater temperature for the depth of migrators (which must be done in the Binning step; [Figure 1](#)). After creating a constant temperature file the day/night code must be run. For our analysis, we chose an average midwater depth temperature of 17°C. The active flux is then assessed by summing the migratory ‘M’ CO₂ column for the nets that are part of your EZ in `Output\PairedNets\YOUR_PAIR\Your_Pair_T17_DayNightNETS.csv`.

An alternate method made possible by whole water column imaging datasets is a direct calculation of the biomass and associated physiological activity of the migratory plankton at the depth to which they migrate during the day. This calculation is made by summing the physiological contributions at their more precise midwater temperature *in situ* in the twilight zone (TZ). This is done by summing the migratory ‘M’ CO₂ column for the nets that are part of your EZ

in `Output\PairedNets\YOUR_PAIR\Your_Pair_DayNightNETS.csv`.

Next, we wanted to test the importance of taxonomic specificity in the pipeline. We ran both EZ and TZ based assessments using our calanoid copepod biovolume to DM conversion equation, and assuming copepod physiological coefficients. We additionally used this output to compare the effect of our taxonomic and non-taxonomic biomass to biovolume calculations by comparing our results with the biomass measurements taken on the BATS cruises before and after the cruise on which our MOCNESS were run. These tows are distinctly different in nature; they are taken with replicate daytime and night time tows every month using a 1-m oblique tow of a 202-μm mesh net to between 150–200 m (depth recorded with a Vemco Minilog recorder) as previously detailed (Madin *et al.*, 2001, Steinberg *et al.*, 2012). As the depth of capture for these tows varies, we used their biomass m^{−2} to our calculated EZ biomass m^{−2}.

Biomass-based methods for active flux have previously been calculated by measuring the total biomass of a size fraction of zooplankton (mg) and either estimating the number of individuals in a tow (counts), or estimating the average size of a single individual (mg). Using these two measurements the allometric equations can be applied to an estimated average sized zooplankton within a size fraction and scaled the estimated number of individuals within that size fraction (Al-Mutairi & Landry, 2001, Kwong *et al.*, 2020, Steinberg *et al.*, 2012). Unlike these prior methods, we have both a direct measurement of the number of individuals and a direct measurement of the biovolume of each individual. Our uncertainty comes, instead, from the biovolume to biomass conversion. We wanted to assess the consequences of assuming an average biomass (with no taxonomy) rather than utilizing the range of biomasses we measured. To do so, we needed to convert our bin sizes to a set of values comparable to the standard mesh sizes for zooplankton size fractionation (200, 500, 1000, 2000, 5000 μm). We did this by assuming that the minor axis of an individual should closely match the sieve

size. The chosen cutoff was biovolume bins of 0.0177, 0.1, 1, 10 and 56.23 mm³, which equates to a minor axis of 270, 457, 922, 1862 and 3154 µm, respectively (based on the minor axis to biomass equation generated using the data provided in [Supplementary File 1](#) where $\text{minor_axis_mm} = 0.922(\text{biovolume_mm}^3)^{0.3052}$; $R^2 = 0.93$). The calculated total biomass (mg) of each of these bins was compared with the measured number of particles to give an estimated dry weight per individual per size class. We then used the sum of the biomass for each size fraction and the average size individual of a size class to calculate the respiratory flux from the EZ using the 17°C midwater temperature and the copepod-specific allometric equations described previously.

RESULTS: USING THE DATA

Biomass to biovolume validation

Biovolume and dry mass were significantly statistically related ($P < 0.001$) for our dataset, with dry mass = 0.029(biovolume). The slope of the relationship varied among taxonomic groups ([Table III](#); [Figure 2](#)), with cyclopoid copepods having a much higher slope and lower fit (R^2). Additionally the groups containing calcifying organisms, such as the foraminifera and the mollusca, which in the region are typified by the shelled atlantiid heteropods and thecosome pteropods, had substantially higher slopes and a lower fit than the remainder of the community. The slope of the total dataset did not, however, change substantially with removal of these taxonomic groups (0.028). Similarly, area and dry mass were significantly related ($P < 0.001$), with dry mass = $0.038(\text{area})^{1.064}$ for the full dataset. Area equations had a higher R^2 for some taxa ([Table III](#)), but biovolume was a better fit for the major migrators (calanoid copepods and euphausiids; [Steinberg et al., 2002](#)) and the aggregate equation.

Binning code output

The output of our initial analysis pipeline (Binning; [Figure 1](#)) results in a set of summary spreadsheets (M#_bins.csv and M#_nets.csv files) for each tow with associated summary plots (M#_BM_bynet.png, M#_O2_bynet.png). The spreadsheets are used in all subsequent calculations and pipelines while the visuals are useful for determining the relative importance of the various size fractions within a net, and setting the

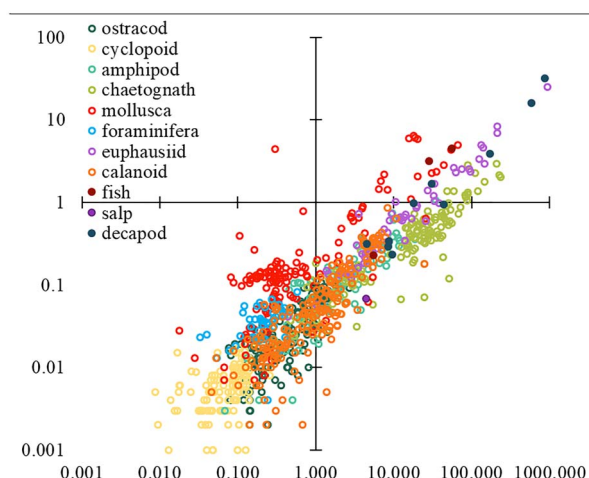


Fig. 2. Biomass to Biovolume dataset. The estimated biovolume of each individual organism (mm³ using an estimated ellipsoidal shape) is highly linearly correlated to the measured dry mass (mg), while the area (mm², calculated using major and minor axis) is strongly correlated as a power function for most taxonomic groups.

limits on x and y -axes for downstream plots. Once the summary files are constructed, the datasets feeding into three additional pipelines.

Waterfall plots (PSD)

The first set of processing code results in waterfall plots of particle size distributions (PSD) with a spreadsheet of the slopes and fit of the linear regression between size (binned biovolume, x -axis) and particle density (m³, y -axis) or abundance (m², y -axis) in the various net strata (output as Your_Tow_Waterfall.png). Analysis of the PSD plots from our dataset suggest that the 150 µm MOCNESS net sampling strategy does not effectively capture organisms under the 0.0177 mm³ size class bin, as these typically pass through the net mesh ([Figure 3](#)). This assessment was made by observing where the smaller bin size had a lower abundance and density than the next larger sized bin. Similarly, our data suggest that particles over and including the 100 mm³ bin are too rare to be adequately sampled in a statistically significant way. This assessment was made quantitatively by observing bins that had, on average, less than five counted particles in a net, which roughly correlated with one individual m². We did this by creating a pivot table in excel using the output from the M#_bins.csv file. We made rows clustered by net and by bin and then looked at the summed counts and the summed NFreq_m3. From these plots, we can conclude that all further calculations are likely only accurate between the 0.0177 and 56.23 mm³ size bins. For all

Table III: Taxon-specific relationship between biovolume and dry mass

Taxa	n	Biovolume (linear)		Area (power)		
		Slope	R ²	Coefficient	Exponent	R ²
Calanoid copepod	183	0.055	0.821	0.027	1.387	0.745
Ostracod	147	0.052	0.748	0.029	1.268	0.680
Mollusca	139	0.097	0.476	0.132	1.019	0.622
Chaetognath	130	0.013	0.732	0.025	1.005	0.827
Cyclopoid copepod	126	0.074	0.022	0.013	0.631	0.175
Amphipod	76	0.034	0.750	0.029	1.296	0.782
Euphausiid	59	0.027	0.978	0.026	1.381	0.967
Foraminifera	43	0.142	0.015	0.048	0.562	0.106
Decapod	10	0.034	0.972	0.020	1.457	0.986
Fish	3					
Salp	1					
All	918	0.029	0.849	0.038	1.064	0.755

The regressions provide the conversion factors by which image data can be transformed into standard weight (mg) using a linear equation with the intercept=0 for biovolume (mm³) or a power equation for area (mm²). The raw data including minor axis, major axis, ESD, area, biovolume, dry weight and wet weight are available in [Supplementary File 1](#).

Table IV: Particle size distribution; the average slope of the PSD for each net and time of day was calculated along with its standard error (n = 4 except * where n = 3)

Depth (m)	Day				Night			
	Slope	SE	R ²	SE	Slope	SE	R ²	SE
Net 8	−1.15	0.086	0.97	0.008	−1.05*	0.066*	0.98*	0.005*
Net 7	−1.30	0.034	0.94	0.016	−1.06	0.042	0.97	0.009
Net 6	−0.94	0.046	0.97	0.009	−0.86	0.064	0.98	0.003
Net 5	−0.83	0.029	0.92	0.021	−0.83	0.059	0.94	0.010
Net 4	−0.72	0.046	0.94	0.003	−0.72	0.050	0.92	0.008
Net 3	−0.59	0.031	0.93	0.006	−0.70	0.026	0.96	0.009
Net 2	−0.64	0.027	0.95	0.011	−0.70	0.050	0.95	0.014
Net 1	−0.63	0.085	0.92	0.006	−0.65	0.053	0.94	0.014

The depth of the nets (m) ranged slightly between cruises (see [Table I](#) for exact depths). A higher slope suggests a greater proportion of small individuals. Also reported is the fit of the line (R²) and its standard error. Lower R² tended to correlate with depths where migratory populations were most abundant. The slopes and statistics for these regressions were drawn from the R output YOUR_PATH\WF_Your_Tow_reg_stats.txt.

subsequent analyses, we implemented this size cut-off in the code.

Using our full dataset (all 8 tows), we observed that within the EZ (0–200 m), there are 2 to 10 times as many individuals (# m²) compared to the TZ depth strata, with a maximum abundance around the chlorophyll max (50–200 m). In these two EZ depth intervals, there are a proportionally greater numbers of small individuals, resulting in a steeper (more negative) ‘waterfall plot’ of size class and abundance ([Figure 3](#); [Table IV](#); the intercepts, slopes and statistics for these regressions were drawn from M#_TowInfo_WF_Your_Tow_reg_stats.txt). This is especially evident during the day when larger migrators move deeper in the water column. The slope of the PSD in the twilight zone (200–1000 m) is lower and tends to decrease with depth ([Table IV](#)). Regressions have a comparatively slightly lower R² during

the day, indicating a less linear profile and suggesting depths where migrators are a significant portion of the community.

In the EZ, there were, on average, 1.5 times as many individuals at night as during the day. This was calculated by comparing the summed abundance (# per m²) for all of the bins in the well sampled range (.01 < bins < 100 mm³) in net 7 and 8 using Tow_bins.csv of a day/night pair, or only net 7 in July 2018 when the surface net was missing. In biovolume, this difference was similar, with 1.4 times as much biovolume during the night. Taxonomically resolved conversion of biovolume to biomass revealed that estimates of total biomass in the EZ range from 0.101 to 0.382 for the day and 0.277–0.794 for the night, resulting in a night:day biomass relationships falling between 1.6 and 3.0. Our non-taxonomically resolved estimates (using the calanoid copepod conversion factor)

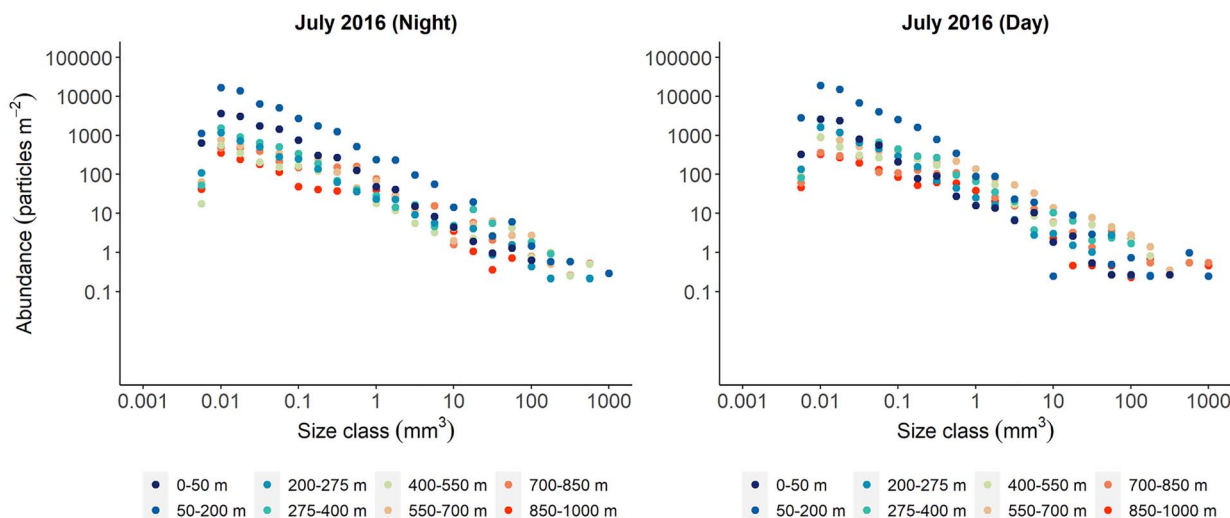


Fig. 3. Your_Tow_Waterfall.png output. This plot depicts the particle abundance (number m^{-2}) based on size class bin (biovolume, mm^3) during the day or the night for each depth region (also sometimes referred to as a particle size distribution; PSD).

were consistently higher by an average of $20 \pm 2\%$ (SE), but retained a very similar night:day biomass relationship falling between 1.8 and 3.1 within the EZ.

Zooplankton daytime net avoidance or patchiness was observable in our sampling, with total biovolume ($\text{mm}^3 \text{ m}^{-2}$ summed 0–1000 m for all pairs except July 2018 where it was summed 50–1000), ranging from a 1.5 to 1.1 times increase in night time tows compared with paired daytime tows. This was calculated by comparing the summed abundance (biovolume m^2) for all of the bins in the well sampled range ($.01 < \text{bins} < 100 \text{mm}^3$) in a day/night pair.

Day/night summaries

The second set of products are the summed biovolume (mm^3), biomass (mg) and physiological contribution (in this case, $\mu\text{mol O}_2 \text{ ind}^{-1} \text{ h}^{-1}$, or $\mu\text{mol CO}_2 \text{ ind}^{-1} \text{ h}^{-1}$) by either m^3 or m^2 . These are reported both by net and by bin size. Most downstream analyses use m^2 as this allows for comparisons among nets of different size intervals. We used these products to make calculations of (i) variations in total biovolume among tows and patchiness/net avoidance within net pairs and (ii) the relative contributions of the migratory versus resident zooplankton community to biomass or apparent oxygen utilization (AOU).

The calculated biomass and AOU contributions of the resident and migratory communities come directly from the spreadsheet output of the Day/Night scripts (using the 0.0177 to 56.23 mm^3 bins). These are saved as Your_Tow_DayNightNETS.csv. They are visualized by the plots saved at Your_Pair_DNPlot.png (Figure 4).

These show the nets plotted on the y -axis with the deepest on the bottom. The summed biomass is the x -axis in panel A, and the summed AOU is the x -axis in panel B. The code could as easily plot summed biovolume or CO_2 if that is the scientific question. The plot shows the non-migratory contributions in black stacked with the migratory contributions in gray, but makes no distinction of the time of day that the migrators are present in the net. To determine whether the migrators are more abundant during the day or the night, a flip plot is provided in panel C. This is the day minus night biomass, resulting in negative numbers being the location and contribution of the sum of the migratory biomass during the night and positive numbers demarcating the day. From our dataset, we can say that resident organisms contribute between 45 and 64% or 44 and 65% of the total biomass (mg DW) 0–1000 m (calculated by comparing—the resident ‘R’ and migratory ‘M’ summed values from Your_Tow_DayNightNETS.csv using both the taxon-specific or generalized biovolume to biomass equations, respectively). Owing to the abundance of smaller individuals in the resident communities, these organisms consequently play a very substantial role in the associated biogeochemical signatures, contributing 56–87% of the total AOU in the water column (using both taxon-specific and generalized equations).

Migration heatmaps

The final section of code focuses on the size specific movement and contributions of the migratory zooplankton. The utility of these outputs is that it allows for

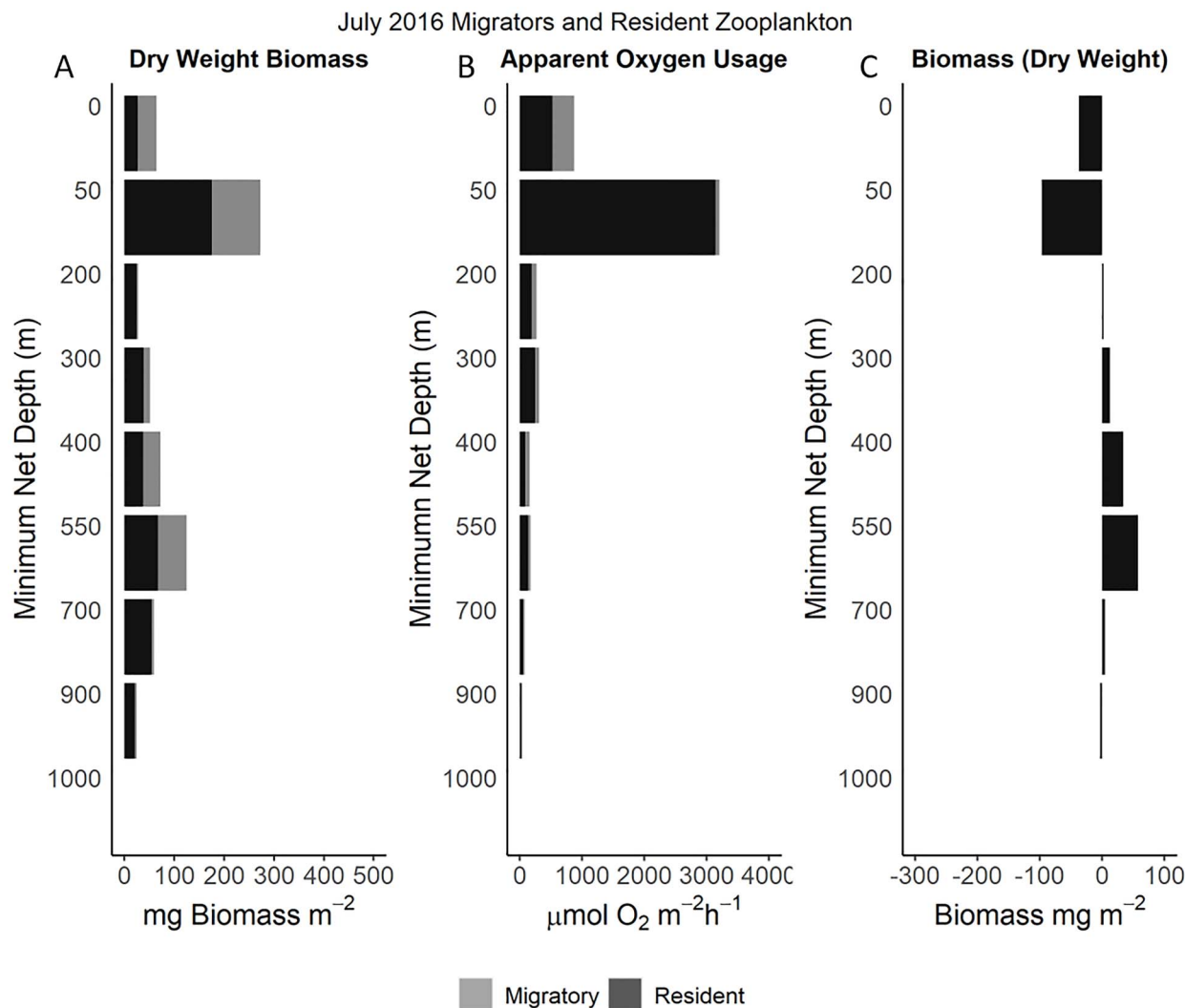


Fig. 4. Your_Tow_DNPlot.png output. The total estimated biomass (**A**) and apparent oxygen usage (**B**) plots calculate residents (black) as the smaller of the day and night values in a net, while the migratory (gray) are the absolute value of the day minus night for the depth strata. The directionality of the migration (**C**) is the raw value of the day minus night for the depth strata. Consequently, negative values demonstrate where migratory organisms are during the night and positive values are where they are during the day.

calculations of (i) the size specific location and extent of the vertical migration (ii) the estimated active flux of the migratory community. The output consists of a.csv with the Your_Pair_DayNightBINS.csv, which provides the day minus night contributions summed within each bin and a heatmap graphical representation of your metric (biovolume, biomass, O₂, CO₂, etc.) as Your_Pair_Metric_Shift.png. The heatmap has the bin size (in our case biovolume mm³, restricted to 0.0177 to 56.23 mm³) on the *x*-axis, net on the *y*-axis (with 0 m at the top), and the colors represent the day minus night of the metric of choice. Red colors represent the summed

contributions to that bin during the daytime and blue represents the contributions during the night time.

Detailed analysis of the depth of migration of the various size bins in our dataset reveals that the region above the local oxygen minimum zone (nets 3 & 4; ~400–700 m) has the bulk of the larger migratory community (>1 mm³), although shorter DVM behaviors are observable in all classes (Figure 6A; saved as Your_Pair_BV_Shift.png). There is a clear pattern of larger individuals migrating to deeper daytime and shallower night time depths. Consequently, there were typically two depths where there were peaks in the

Table V: Calculations of respiratory CO₂ production (mmol m⁻² d⁻¹)

	Taxonomic			Non-taxonomic			
	17° (EZ)	17° (TZ)	IS (TZ)	17° (EZ)	17° (TZ)	IS (TZ)	Avg. size 17° (EZ)
July 2016	2.58	1.8	1.91	2.49	1.76	1.87	0.26
July 2017	16.69	0.52	0.54	16.41	0.54	0.55	0.83
October 2018	9.26	1.38	1.78	9.16	1.38	1.78	0.41

Migratory organisms were characterized as day minus night for each size bin for paired tows. The first estimate was made by calculating the physiology of each individual particle from 0 to 200 (euphotic zone; EZ) at midwater temperatures (17°C). Next the active flux was calculated based on adding together the respiratory CO₂ production of all migrants present during the day in the 200–1000 m (twilight zone; TZ) region, calculating using both a standard 17°C midwater temperature and in-situ (IS) temperatures. The final analysis uses an averaged size individual per size class (See Table VI) and applies this to the EZ biovolume calculations at a standard 17°C midwater temperature. All analyses assume a 12 h day.

migrator biomass and associated deposition of active flux materials into the water column (Figure 6B; saved as Your_Pair_CO2_Shift.png; Supplementary File 3 for all plots). The first was right below the euphotic zone (net 6; ~200–300 m) where excreta material from numerous smaller individuals (<1 mm³) would be delivered. Additional flux would be delivered to greater depths by a large size range of organisms, consistently peaking in nets 3 to 5 (between ~300–700 m). Additionally, there is potential evidence of deeper organisms of a range of size classes migrating up into the deepest nets during the night by the blue shading in the lowest depths (i.e. Vinogradov's ladder; Vinogradov, 1962). This could, alternatively, be a consequence of individuals migrating down during the day to avoid the influx of larger migrators from the surface.

Active flux calculations

For the three paired day/night tows where all nets were available for active flux calculations, the particle-based EZ estimates of active flux were consistently greater than TZ calculations. When the integrated biomass 0–1000 was substantially different between paired day/night tows (July 2017 and October 2018), the EZ flux was substantially higher than TZ estimates, ranging from a factor of 1.5 to 30 times greater flux (Table V). The choice of midwater temperature minimally effected flux calculations within the TZ; using *in situ* temperatures increased the active flux calculation between 2 and 29% (Table V). Use of non-taxonomic coefficients and conversion factors (i.e. treating every particle as if it were a calanoid copepod) had an even less profound effect, ranging between 1 and 4% variation between methods, with no consistent directionality in estimation error.

Binning of individuals within the standard size fractions demonstrated both diel and vertical variation in the average size of zooplankton, particularly in the larger size class (Table VI). Assumption of an average particle size

for the day and night throughout the EZ resulted in a 1.5 times difference in the number of particles between methods, with no consistent directionality and the greatest variability in the 2000 µm size class. This was calculated by comparing the summed measured number of particles in a net (from Your_Tow_bins.csv), and the number of particles calculated by dividing the total biomass in a net (the summed biomass from Your_Tow_bins.csv) by an average sized individual. Although the variation in the particle size of the smallest fraction was small, the higher number of particles in the smallest size fractions and the allometric relationship associated with metabolic function, resulted in greater impacts on the calculation of active flux in the smallest size fractions, causing large discrepancies between the average biomass method and the image based method. As a consequence, the total flux is greatly and consistently underestimated when using average sizes compared to using individual particles (10 to 23 times lower; comparing EZ calculations at 17°C; Table V).

DISCUSSION

Application of a set of R codes to a size-based dataset of the Sargasso Sea zooplankton community demonstrates the usefulness of size-based datasets to quantitatively describe allometrically constrained ecological characteristics of the zooplankton community. The analytical pipeline allows rapid and consistent handling of large image-based datasets, with the ability to apply taxon-specific coding of various allometric and biomass equations. Application of the code to a Sargasso Sea depth-stratified dataset emphasizes the dominance of resident zooplankton in the ecology of midwater communities in the region, provides valuable information about the variability in size structure of the euphotic and twilight zones over the diel cycle, and suggests important and somewhat

Table VI: Average particle dry mass variation

Fraction	Average particle mass (mg)								
	Avg.	0–50 m				50–200 m			
		Day		Night*		Day		Night	
200 μm	0.0021	0.0021	± 0.0001	0.0021	± 0.0001	0.0020	± 0.0000	0.0021	± 0.0000
500 μm	0.0140	0.0134	± 0.0009	0.0142	± 0.0001	0.0138	± 0.0007	0.0147	± 0.0001
1 000 μm	0.1473	0.1398	± 0.0209	0.1397	± 0.0128	0.1593	± 0.0146	0.1486	± 0.0089
2000 μm	1.3120	1.1810	± 0.2240	1.5128	± 0.1314	1.3249	± 0.1861	1.2797	± 0.1160

We compared the estimated average zooplankton mass based on the full PSD and counts using taxonomic dry mass conversion factors (average $\pm$ SE; $n=4$ unless * where $n=3$). An average mass for the EZ was also calculated (Avg.) for analysis of the effect of particle mass on active flux calculations (Table V).

consistent size-based structure in vertical migration patterns. It also highlights net sampling inefficiency, as well as the limitations of day–night estimation of vertical migration of the upper water column (EZ).

In our PSD dataset, the presence of migrators in the midwater (TZ) stood out as peaks above the average slope of the line or below the average slope of the line in the photic zone (EZ) during the daytime, resulting in a departure from linearity and slightly lower R^2 statistics. In contrast, the slope of all of the PSD had the highest R^2 across most of the depth strata during the night when the migrators were at the surface. This implies that the migratory community is ecologically and energetically more closely tied to the abundance and community size structure of the EZ, rather than the TZ. Experimentally derived data of this type will be valuable to begin to test hypotheses about functional size spectrum models of zooplankton community structure (Heneghan *et al.*, 2020).

Previous estimates of zooplankton biomass in the region document a daytime average of 0.418 g m^{-2} and a nighttime value of 0.659 g m^{-2} and suggest that July is a period of higher biomass than October (Madin *et al.*, 2001, Steinberg *et al.*, 2012). This is consistent with our dataset, whose taxonomically resolved estimates of biomass from biovolume measurements range from 0.101 to 0.382 for the day and 0.277 to 0.794 for the night. The smaller mesh size of our nets (150 vs 202 μm) would likely be far outweighed by the fact that we cut our sampling at $>10 \text{ mm}^3$ biovolume while prior analyses included all larger taxa. When our estimates are compared directly with the weighed biomass from BATS cruises taken before and after our sampling time points, both the taxonomically and non-taxonomically resolved biomass estimates fell close to the measured values, with no apparent better estimate, although non-taxonomically resolved estimates were consistently on average 20% higher than those made with taxonomic equations (Figure 5). As the BATS samples and our MOCNESS tows occurred anywhere between 1 and

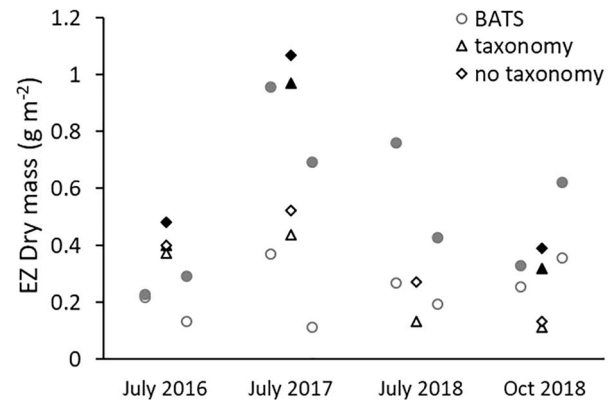


Fig. 5. Biomass comparison. We directly compared the EZ biomass data estimated from our biovolume calculations with measured zooplankton biomass from the BATS (gray circles) time series cruises that happened 1–3 weeks before and after our sampling. Daytime (open shapes) and night time (closed shapes) are on a similar order of magnitude to both our taxonomically resolved (black triangles) and non-taxonomically resolved (black diamond) estimates.

3 weeks apart, and with some spatial heterogeneity, it is likely that small-scale patchiness and mesoscale features contribute to our inability to determine which biomass equations are regionally more appropriate. As our values fell within the range directly measured at BATS, it does give confidence that the method provides a reasonable approximation of the community biomass. Based on these estimates, our night:day biomass relationships fall between 1.6 and 3.1, which is also consistent with previous data (average 1.2; range = 0.3–12.3; Madin *et al.*, 2001, Steinberg *et al.*, 2012).

An interesting aspect of this analysis is that it demonstrates that resident (i.e. non-migratory) organisms play a larger role in the supply of dissolved nutrients and demand of O_2 than the DVM community in the midwater. These residents ranged from 30 to 60% of the total biomass in the EZ, but were 60 to 78% of the total biomass in the TZ. As such, midwater processes

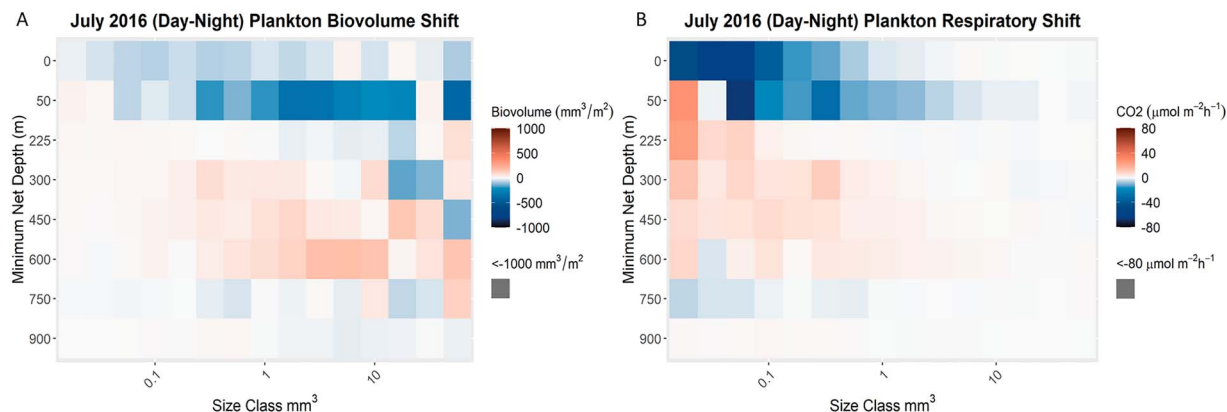


Fig. 6. Your_Tow_Shift.png output. These figures depict the day minus night biomass ($\text{mm}^3 \text{m}^{-2}$) and day minus night CO_2 active flux ($\mu\text{mol m}^{-2} \text{h}^{-1}$) from each size class and net from MOCNESS day/night pairs. Bins (in our case mm^3) are on the x -axis and nets are arranged from shallow to deep on the y -axis. Color demarcates the sum of the day minus night of the metric for a specific size bin and net. Positive (warm) numbers indicate increased biomass or flux during the day while negative (cool) numbers indicate increased biomass or flux during the night.

and interactions are strongly driven by the resident community, for which very little is known (Steinberg *et al.*, 2008a). These organisms, whose food source is derived from surface carbon export, are detritivores on particulate organic carbon (POC), or carnivores, consuming migrators and each other. These midwater species are thus associated with attenuation of flux from the surface (Stukel *et al.*, 2019). The food web dynamics, abundance, biodiversity and ecology of these ‘dark ocean’ organisms is one of the most poorly constrained components of the biological pump (Burd *et al.*, 2010). Further classification to large taxonomic groups (e.g. copepod, pteropod, euphausiid, etc.) would increase our understanding of which specific groups contribute to midwater processes and vertical migration and potentially provide insight into their trophic role and is possible by implementing the taxonomic filters provided in the R code.

Analysis of size specific vertical migration patterns demonstrates how size influences the depth of the observed migration. Although the total biomass shifted between seasons and years, the depth of daytime abundance typically increased with increasing biovolume. Similarly, the night time depth was occasionally deeper for larger migratory organisms. This is consistent with the observations from Ohman and Romagnan (2016), who reported size specific DVM patterns in copepods from the California Current System. As demonstrated by our TZ assessment of active flux, as well as the model results of Archibald *et al.* (2019), the specific depth of migration of an organism is important for estimating its active flux. Broader geographic and finer taxonomic assessment of this size-based depth distribution will be

valuable to determine how phylogenetic patterns interact with hydrographic features, such as light attenuation and midwater oxygen, to provide predictive equations for the depth to which specific size class zooplankton migrate during both the day and night. Detailed assessment will be especially important in regions with oxygen minimum zones, which clearly influence both the distribution and the midwater physiology of organism (Kiko & Hauss, 2019, Seibel, 2011, Wishner *et al.*, 2018). Incorporating this response to low oxygen into our allometric equations will be critical to making accurate assessments of biogeochemical fluxes.

Using the output of our code we chose to explore four sources of variation in the calculation of active flux (i) the choice of the day minus night calculation from the EZ versus the TZ, (ii) the choice of a constant versus in situ-based midwater temperature (iii) the use of taxonomically specific coefficients and conversion factors and (iv) the choice of actual versus average size individuals. The objective of this analysis was to demonstrate both why optically-based depth stratified methods are an improvement and to try to make comparisons with prior biomass-based approaches. As it is well known, the larger nektonic size classes, which were represented by the size bins greater than 100 mm^3 ($\sim 2.9 \text{ mg DM}$; corresponding to $> 3.75 \text{ mm}$ minor axis particles), were poorly sample by the $150 \mu\text{m}$ nets, likely both due to natural patchiness and their greater net avoidance (Fleminger & Clutter, 1965, Ianson *et al.*, 2004, Skjoldal *et al.*, 2013). Although both patchiness and net avoidance could explain differences between the EZ and TZ calculations, the consistently higher 0–1000 m biomass during the night time suggest

that net avoidance in the EZ is a major contributor. This is problematic as the biomass from these top nets are historically those used to calculate the active flux. Consequently, our active flux measurements made from the EZ were highly variable (2.6–16.7 mmol CO₂ using taxon-specific results) and up to 32 times greater than their paired TZ measurements. In contrast, when we calculated active flux based exclusively on the migratory biomass 200–1000 m (TZ estimates), we consistently calculated less variable respiratory flux (0.52–1.9 mmol CO₂, irrespective of the temperature choice using taxon-specific results). It is important to note that the calculation made using the summed TZ physiology may be influenced by the migration of individuals up into the TZ from below.

The biomass to biovolume and area conversion factors documented herein provide the first such analysis for ZooScan data. There is a long history in the literature linking various optical methods and biomass or compositional characters using photography (Alcaraz *et al.*, 2003, Davis & Wiebe, 1985, Lehetto & Hernández-León, 2009); however, each method and region tends to result in slightly different regression equations. Compared to these earlier works, our regressions tended to have lower R^2 . For example, the length to dry mass equations from Davis and Wiebe (1985) have an R^2 almost exclusively above 0.9, while the area to dry mass equations for various aggregate crustacean assemblages from Lehetto and Hernández-León (2009) all fall above an R^2 of 0.9, although their monospecific and non-crustacean equations had R^2 of, on average 0.72 and a range from 0.26 to 0.94. The exponents of our area to biomass equations (0.63–1.45) were similar to that of the Lehetto dataset (0.63–1.78), with higher exponents for crustacean zooplankton. The coefficients in the Lehetto dataset were, however, consistently greater than in our dataset (with crustaceans averaging at 0.045 compared to our 0.026). These differences may result from regional or taxonomic variation, as the Lehetto dataset was generated by analyzing upper water column individuals from the Canary Islands (0–200 m) and Antarctica (0–400 m), where individuals skewed larger than in our dataset. Additionally, there are known effects of preservation on biovolume and dry mass (Ahlstrom & Thrailkill, 1963, Böttger-Schnack & Schnack, 1986). While our samples were overwhelming preserved in formalin, those from Lehetto were imaged fresh. Additionally, it is possible that methodological differences in machine versus human measurement contribute. In the ZooScan method, the computer may pick major and minor axes differently than in the photography or silhouette method, by which a human demarcates the axes of an individual and would likely incorporate more morphological understanding of the image.

Inclusion of taxon-specific biomass to biovolume conversions and coefficients in the allometric equations resulted in, on average, a 2% variation in active flux estimates. This is partially due to the fact that the community was very strongly dominated by calanoid copepods (typically close to 50%), so choosing their constants as the default code well approximated the community composition. An additional consideration for the biovolume equations is the choice of the intercept. Our initial observations were that it was important to set the intercept to 0, as otherwise the cumulative effect of numerous small particles became problematic (especially when the intercept was negative). For the neutrally buoyant organisms this seems like a philosophically reasonable decision. For the foraminifera and shelled mollusca (heteropods and thecosomes pteropods), it may be logical to allow a positive intercept. The ZooScan image to biomass conversion equations from this dataset are thus valuable for converting datasets into biogeochemical estimates but will require further validation in a regionally more diverse dataset.

Our exploration of the temperature effect on physiological calculations demonstrated that variation in the active flux calculations due to an assumption of a static depth of migration (and associated static temperature) for their Q_{10} corrections is relatively small. The *in situ* temperatures estimated 2–29% more flux using the taxon-specific dataset, which is insufficient to account for the discrepancies between EZ and TZ assessments (which was minimally 140% different). This temperature correction may, however, be more problematic in regions with a steeper temperature decline with depth. These findings support the conclusion that calculations using the TZ approach would resolve difficulties with net avoidance during the daytime, which has been modeled to overestimate active flux by 50% in the Azores Front region (Iansson *et al.*, 2004). Although midwater (TZ) measurements of active flux are substantially more time consuming and require specialized equipment (i.e. MOCNESS and ZooSCAN) to conduct, their insensitivity to net avoidance makes them preferable to the more standard EZ assessment.

Prior estimates of active flux at the BATS site used biomass data from the EZ, assumed an average size, and estimated a midwater temperature of 18°C (Steinberg *et al.*, 2000) or 17°C (Steinberg *et al.*, 2012). These assessments found respiratory CO₂ values with a mean of 0.13 (max 0.53) or of 0.21 (± 0.07 SD) mmol m⁻² d⁻¹, respectively. These values are in line with our analyses that use similar methods: i.e. the non-taxonomically resolved EZ estimates based on average sized particles at 17°C (0.26, 0.41 and 0.83 mmol m⁻² d⁻¹). Comparing the calculations based on an average sized individual versus the full size spectra of particles for analysis of the active

flux from the EZ demonstrated that variations in the average size of a particle within a size class resulted in substantially higher (10 to 23 times) total active flux estimates. This discrepancy comes from the highly skewed size distribution, with the power relationship between organismal size and abundance causing small individuals to be numerically dominant in each size fraction. In addition to their greater abundance, the higher mass specific physiology of smaller organisms (also based on a power function) causes their contributions to play an outsized role in the estimate of flux.

From our observation, we can say that our pipeline seems to yield similar results to older biomass-based methods when making the same assumptions, but additionally indicates that many of those assumptions substantially underestimate the true active flux contributions of the migratory community. Our results indicate that the most critical factor for accurate active flux estimates is precise characterization of organismal size/abundance (consistently >1000% difference). The next is the use of a midwater assessment of migrator biomass, as it removes problems of net avoidance (at least >150% difference, but highly variable). Assignment of a precise midwater temperature correction is the next most important methodological choice (2–29% difference), while application of taxonomically resolved conversion factors is the least important (2% difference, in this region).

Although the method presented here provides useful quantitative tools to describe the distributional and migrational patterns of the zooplankton, it is substantially more laborious and expensive than previously used net sampling protocols. The biogeochemical conversion factors used in this assessment, particularly the biovolume to biomass equations, will need to be tested in multiple locations and with communities from a range of depths. Most experimentally derived allometric equations for physiological data are heavily skewed to represent crustacean groups. At BATS, this is not a problem as the community is dominated by crustacea (Blanco-Bercial, 2020, Ivory *et al.*, 2019), but that assumption would not hold in all regions or seasons.

Additionally, most physiological measurements have been made with either active migrators or surface species, and the estimates of the physiological contributions of midwater organisms may need to be corrected for a less active mode of life (Seibel & Drazen, 2007). Applying a depth correction based on depth of capture for the physiology of ‘resident’ plankton using the equations provided in Ikeda (2014b), would be one approach to account for this proposed slower physiology at depth. The actual physiological rates of most midwater organisms have not, however, been well validated although there are ongoing efforts to better resolve the midwater metabolism

(Hernández-León *et al.*, 2019). Better predictions of allometric relationships across taxa and habitats should be a research priority to make imaged based datasets better predictors of zooplankton contributions to biogeochemical cycles. Despite the limitations associated with the depth-stratified net sampling and the embedded physiological assumptions, this allometric analysis pipeline provides a framework for quantitatively comparing and testing hypotheses about the distribution, migration patterns and biogeochemical impacts of mesozooplankton.

CONCLUSIONS

We developed an R code that allows for a consistent and quantitative analysis of mesozooplankton particle size distribution, total biomass and physiological contributions of migratory versus resident organisms, size-based variation in the depth of migration, and the calculation of active flux directly from Ecotaxa output. The code can be modified to include any allometric equation for biogeochemistry and is capable of implementing taxonomic specific coefficients or filters. Application of this code on a set of paired day/night MOCNESS tows 0–1000 m from the Sargasso Sea near BATS demonstrate the importance of resident zooplankton in the ecology of plankton communities and the importance of migrator size on the depth of migration. Calculations of active flux resulting from this pipeline reveal the importance of the smaller size class organisms and emphasize the improvements of depth stratified size-based assessments of zooplankton biogeochemical contributions.

SUPPLEMENTARY DATA

Supplementary data can be found at *Journal of Plankton Research* online.

ACKNOWLEDGEMENTS

Cruise and at sea support was provided by the captains and crew of the *R/V Atlantic Explorer*, and sampling was with the assistance of the broader BIOS-SCOPE team. We are grateful for the assistance of Joe Cope and Joe Jones for taking the July 2017 samples on our behalf. Thanks to Andrea Miccoli for help at sea, preliminary data processing and scanning.

FUNDING

Funding for this project was provided by Simon's Foundation International's as part of BIOS-SCOPE to L.B.B. and A.E.M., National Aeronautics and Space Administration [EXPORTS grant #80NSSC17K0654] to A.E.M., National Science Foundation [OCE grant #1829318] to L.B.B. and A.E.M., and NSF [OCE grant #1948162] to L.B.B. and A.E.M. We appreciate the ship and wire time provided by the Oceanic Flux Program [National Science Foundation

OCE grant #1829885] for our October 2018 sampling time point, and the zooplankton dataset from the Bermuda Atlantic Time Series [National Science Foundation OCE grant #1258622 and #1756105].

DATA ARCHIVING

The taxonomy, optical measurements and biomass dataset are available in [Supplementary file 1](#). R-code is available in [Supplementary file 2](#). Particle size data, the output from Ecotaxa used in this analysis, are available along with the R code at <https://github.com/blancoercial/BiogeochemicalContributions>. These data are also publically available in BCO-DMO associated with project 817 889.

REFERENCES

- Ahlstrom, E. H. and Thraillkill, J. R. (1963) Plankton volume loss with time of preservation. *Calif. Coop. Oceanic Fish. Invest. Rep.*, **9**, 57–73.
- Al-Mutairi, H. and Landry, M. R. (2001) Active export of carbon and nitrogen at station ALOHA by diel migrant zooplankton. *Deep-Sea Res. II*, **48**, 2083–2103.
- Alcaraz, M., Saiz, E., Calbet, A., Trepas, I. and Broglio, E. (2003) Estimating zooplankton biomass through image analysis. *Mar. Biol.*, **143**, 307–315.
- Archibald, K. M., Siegel, D. A. and Doney, S. C. (2019) Modeling the impact of zooplankton diel vertical migration on the carbon export flux of the biological pump. *Global Biogeochem. Cycles*, **33**, 181–199.
- Bianchi, D., Stock, C., Galbraith, E. D. and Sarmiento, J. L. (2013) Diel vertical migration: ecological controls and impacts on the biological pump in a one-dimensional ocean model. *Global Biogeochem. Cycles*, **27**, 478–491.
- Blanco-Bercial, L. (2020) Metabarcoding analyses and seasonality of the zooplankton community at BATS. *Front. Mar. Sci.*, **7**, 173. doi: [10.3389/fmars.2020.00173](https://doi.org/10.3389/fmars.2020.00173).
- Böttger-Schnack, R. and Schnack, D. (1986) On the effect of formaldehyde fixation on the dry weight of copepods. *Meeresforschung = Reports on Marine Research. Berichte der Deutschen Wissenschaftlichen Kommission für Meeresforschung*, **31**, 141–152.
- Burd, A. B., Hansell, D. A., Steinberg, D. K., Anderson, T. R., Aristegui, J., Baltar, F., Beupré, S. R., Buesseler, K. O. *et al.* (2010) Assessing the apparent imbalance between geochemical and biochemical indicators of meso- and bathypelagic biological activity: what the @\$#! Is wrong with present calculations of carbon budgets? *Deep-Sea Res. II*, **57**, 1557–1571.
- Dam, H. G., Roman, M. R. and Youngbluth, M. J. (1995) Downward export of respiratory carbon and dissolved inorganic nitrogen by diel-migrant mesozooplankton at the JGOFS Bermuda time-series station. *Deep-Sea Res. I*, **42**, 1187–1197.
- Davis, C. S. and Wiebe, P. H. (1985) Macrozooplankton biomass in a warm-core Gulf stream ring: time series changes in size structure, taxonomic composition, and vertical distribution. *J. Geophys. Res. Oceans*, **90**, 8871–8884.
- Ducklow, H. W., Steinberg, D. K. and Buesseler, K. O. (2001) Upper Ocean carbon export and the biological pump. *Oceanography*, **14**, 50–58.
- Fleminger, A. and Clutter, R. I. (1965) Avoidance of towed nets by zooplankton I. *Limnol. Oceanogr.*, **10**, 96–104.
- Forest, A., Stemmann, L., Picheral, M., Burdorf, L., Robert, D., Fortier, L. and Babin, M. (2012) Size distribution of particles and zooplankton across the shelf-basin system in Southeast Beaufort Sea: combined results from an underwater vision profiler and vertical net tows. *Biogeosciences*, **9**, 1301–1320.
- Gorsky, G., Ohman, M. D., Picheral, M., Gasparini, S., Stemmann, L., Romagnan, J.-B., Cawood, A., Pesant, S. *et al.* (2010) Digital zooplankton image analysis using the ZooScan integrated system. *J. Plankton Res.*, **32**, 285–303.
- Hansen, P. J., Bjørnsen, P. K. and Hansen, B. W. (1997) Zooplankton grazing and growth: scaling within the 2–2,000-µm body size range. *Limnol. Oceanogr.*, **42**, 687–704.
- Heneghan, R. F., Everett, J. D., Sykes, P., Batten, S. D., Edwards, M., Takahashi, K., Suthers, I. M., Blanchard, J. L. *et al.* (2020) A functional size-spectrum model of the global marine ecosystem that resolves zooplankton composition. *Ecol. Model.*, **435**, 109265.
- Hernández-León, S., Calles, S. and De Puelles, M. L. F. (2019) The estimation of metabolism in the mesopelagic zone: disentangling deep-sea zooplankton respiration. *Prog. Oceanogr.*, **178**, 102163.
- Hirst, A. and Kjørboe, T. (2002) Mortality of marine planktonic copepods: global rates and patterns. *Mar. Ecol. Prog. Ser.*, **230**, 195–209.
- Ianson, D., Jackson, G. A., Angel, M. V., Lampitt, R. S. and Burd, A. B. (2004) Effect of net avoidance on estimates of diel vertical migration. *Limnol. Oceanogr.*, **49**, 2297–2303.
- Ikeda, T. (2013a) Metabolism and chemical composition of marine pelagic amphipods: synthesis toward a global bathymetric model. *J. Oceanogr.*, **69**, 339–355.
- Ikeda, T. (2013b) Respiration and ammonia excretion of euphausiid crustaceans: synthesis toward a global-bathymetric model. *Mar. Biol.*, **160**, 251–262.
- Ikeda, T. (2014a) Metabolism and chemical composition of marine pelagic gastropod molluscs: a synthesis. *J. Oceanogr.*, **70**, 289–305.
- Ikeda, T. (2014b) Respiration and ammonia excretion by marine metazooplankton taxa: synthesis toward a global-bathymetric model. *Mar. Biol.*, **161**, 2753–2766.
- Ikeda, T., Kanno, Y., Ozaki, K. and Shinada, A. (2001) Metabolic rates of epipelagic marine copepods as a function of body mass and temperature. *Mar. Biol.*, **139**, 587–596.
- Ikeda, T. and Takahashi, T. (2012) Synthesis towards a global-bathymetric model of metabolism and chemical composition of marine pelagic chaetognaths. *J. Exp. Mar. Biol. Ecol.*, **424**, 78–88.
- Ivory, J. A., Steinberg, D. K. and Latour, R. J. (2019) Diel, seasonal, and interannual patterns in mesozooplankton abundance in the Sargasso Sea. *ICES J. of Marine Science*, **76**, 217–231.
- Kiko, R., Brandt, P., Christiansen, S., Faustmann, J., Kriest, I., Rodrigues, E., Schütte, F. and Hauss, H. (2020) Zooplankton-mediated fluxes in the eastern tropical North Atlantic. *Front. Mar. Sci.*, **7**, 358.
- Kiko, R. and Hauss, H. (2019) On the estimation of zooplankton-mediated active fluxes in oxygen minimum zone regions. *Front. Mar. Sci.*, **6**, 741.
- Kobari, T., Kitamura, M., Minowa, M., Isami, H., Akamatsu, H., Kawakami, H., Matsumoto, K., Wakita, M. *et al.* (2013) Impacts of the wintertime mesozooplankton community to downward carbon flux in the subarctic and subtropical Pacific oceans. *Deep-Sea Res. I*, **81**, 78–88.
- Kobari, T., Steinberg, D. K., Ueda, A., Tsuda, A., Silver, M. W. and Kitamura, M. (2008) Impacts of ontogenetically migrating copepods on downward carbon flux in the western subarctic Pacific Ocean. *Deep-Sea Res. II*, **55**, 1648–1660.
- Kwong, L. E., Henschke, N., Pakhomov, E. A., Everett, J. D. and Suthers, I. M. (2020) Mesozooplankton and micronekton active carbon transport in contrasting eddies. *Front. Mar. Sci.*, **6**, 825.

- Lehette, P. and Hernandez-Leon, S. (2009) Zooplankton biomass estimation from digitized images: a comparison between subtropical and Antarctic organisms. *Limnol. Oceanogr. Methods*, **7**, 304–308.
- Lombard, F., Boss, E., Waite, A. M., Vogt, M., Uitz, J., Stemann, L., Sosik, H. M., Schulz, J. et al. (2019) Globally consistent quantitative observations of planktonic ecosystems. *Front. Mar. Sci.*, **6**, 196.
- Longhurst, A., Bedo, A., Harrison, W., Head, E. and Sameoto, D. (1990) Vertical flux of respiratory carbon by oceanic diel migrant biota. *Deep Sea Research Part A. Oceanographic Research Papers*, **37**, 685–694.
- Longhurst, A. R. and Harrison, W. G. (1988) Vertical nitrogen flux from the oceanic photic zone by diel migrant zooplankton and nekton. *Deep-Sea Res. I*, **35**, 881–889.
- Longhurst, A. R. and Harrison, W. G. (1989) The biological pump: profiles of plankton production and consumption in the upper ocean. *Prog. Oceanogr.*, **22**, 47–123.
- Maas, A. E., Frazar, S. L., Outram, D. M., Seibel, B. A. and Wishner, K. F. (2014) Fine-scale vertical distribution of macroplankton and micronekton in the eastern tropical North Pacific in association with an oxygen minimum zone. *J. Plankton Res.*, **36**, 1557–1575.
- Maas, A. E., Liu, S., Bolaños, L. M., Widner, B., Parsons, R., Kujawinski, E. B., Blanco-Bercial, L. and Carlson, C. A. (2020) Migratory zooplankton excreta and its influence on prokaryotic communities. *Front. Mar. Sci.*, **7**, 573268. doi: [10.3389/fmars.2020.573268](https://doi.org/10.3389/fmars.2020.573268).
- Maas, A. E., Miccoli, A., Stamieszkin, K. S., Carlson, C. A. and Steinberg, D. K. (2021) Allometry and the calculation of zooplankton metabolism in the subarctic Northeast Pacific Ocean. *J. Plankton Res.* doi: [10.1093/plankt/fbab026](https://doi.org/10.1093/plankt/fbab026).
- Madin, L. P., Horgan, E. F. and Steinberg, D. K. (2001) Zooplankton at the Bermuda Atlantic time-series study (BATS) station: diel, seasonal and interannual variation in biomass, 1994–1998. *Deep-Sea Res. II Top. Stud. Oceanogr.*, **48**, 2063–2082.
- Mayzaud, P., Boutoute, M., Gasparini, S., Mousseau, L. and Lefevre, D. (2005) Respiration in marine zooplankton—the other side of the coin: CO₂. *Limnol. Oceanogr.*, **50**, 291–298.
- Motoda, S. (1959) Devices of simple plankton apparatus. *Memoirs of the Faculty of Fisheries Hokkaido University*, **7**, 73–94.
- Noyon, M., Rasolofarijao, Z., Huggett, J., Ternon, J.-F. and Roberts, M. (2020) Comparison of mesozooplankton communities at three shallow seamounts in the south West Indian Ocean. *Deep-Sea Res. II Top. Stud. Oceanogr.*, **176**, 104759.
- Ohman, M. D., Davis, R. E., Sherman, J. T., Grindley, K. R., Whitmore, B. M., Nickels, C. F. and Ellen, J. S. (2019) Zooglider: an autonomous vehicle for optical and acoustic sensing of zooplankton. *Limnol. Oceanogr. Methods*, **17**, 69–86.
- Ohman, M. D. and Romagnan, J.-B. (2016) Nonlinear effects of body size and optical attenuation on diel vertical migration by zooplankton. *Limnol. Oceanogr.*, **61**, 765–770.
- Picheral, M., Colin, S. and Irissou, J.-O. (2017) *EcoTaxa, a Tool for the Taxonomic Classification of Images*. <http://ecotaxa.obs-vlfr.fr>.
- Robinson, C., Steinberg, D. K., Anderson, T. R., Aristegui, J., Carlson, C. A., Frost, J. R., Ghiglione, J. F., Hernández-León, S. et al. (2010) Mesopelagic zone ecology and biogeochemistry—a synthesis. *Deep-Sea Res. II Top. Stud. Oceanogr.*, **57**, 1504–1518.
- Schnetzer, A. and Steinberg, D. K. (2002) Active transport of particulate organic carbon and nitrogen by vertically migrating zooplankton in the Sargasso Sea. *Mar. Ecol. Prog. Ser.*, **234**, 71–84.
- Seibel, B. A. (2011) Critical oxygen levels and metabolic suppression in oceanic oxygen minimum zones. *J. Exp. Biol.*, **214**, 326–336.
- Seibel, B. A. and Drazen, J. C. (2007) The rate of metabolism in marine animals: environmental constraints, ecological demands and energetic opportunities. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, **362**, 2061–2078.
- Skjoldal, H. R., Wiebe, P. H., Postel, L., Knutsen, T., Kaartvedt, S. and Sameoto, D. D. (2013) Intercomparison of zooplankton (net) sampling systems: results from the ICES/GLOBEC Sea-going workshop. *Prog. Oceanogr.*, **108**, 1–42.
- Steinberg, D. K., Carlson, C. A., Bates, N. R., Goldthwait, S. A., Madin, L. P. and Michaels, A. F. (2000) Zooplankton vertical migration and the active transport of dissolved organic and inorganic carbon in the Sargasso Sea. *Deep Sea Research Part I*, **47**, 137–158.
- Steinberg, D. K., Cope, J. S., Wilson, S. E. and Kobari, T. (2008a) A comparison of mesopelagic mesozooplankton community structure in the subtropical and subarctic North Pacific Ocean. *Deep-Sea Res. II*, **55**, 1615–1635.
- Steinberg, D. K., Goldthwait, S. A. and Hansell, D. A. (2002) Zooplankton vertical migration and the active transport of dissolved organic and inorganic nitrogen in the Sargasso Sea. *Deep-Sea Res. I Oceanogr. Res. Pap.*, **49**, 1445–1461.
- Steinberg, D. K. and Landry, M. R. (2017) Zooplankton and the ocean carbon cycle. *Ann. Rev. Mar. Sci.*, **9**, 413–444.
- Steinberg, D. K., Lomas, M. W. and Cope, J. S. (2012) Long-term increase in mesozooplankton biomass in the Sargasso Sea: linkage to climate and implications for food web dynamics and biogeochemical cycling. *Global Biogeochem. Cycles*, **26**, GB1004. doi: [10.1029/2010GB004026](https://doi.org/10.1029/2010GB004026).
- Steinberg, D. K., Van Mooy, B. A. S., Buesseler, K. O., Boyd, P. W., Kobari, T. and Karl, D. M. (2008b) Bacterial vs. zooplankton control of sinking particle flux in the ocean's twilight zone. *Limnol. Oceanogr.*, **53**, 1327–1338.
- Straile, D. (1997) Gross growth efficiencies of protozoan and metazoan zooplankton and their dependence on food concentration, predator-prey weight ratio, and taxonomic group. *Limnol. Oceanogr.*, **42**, 1375–1385.
- Stukel, M. R., Ohman, M. D., Kelly, T. B. and Biard, T. (2019) The roles of suspension-feeding and flux-feeding zooplankton as gatekeepers of particle flux into the mesopelagic ocean in the Northeast Pacific. *Front. Mar. Sci.*, **6**, 397.
- Vandromme, P., Stemann, L., Garcia-Comas, C., Berline, L., Sun, X. and Gorsky, G. (2012) Assessing biases in computing size spectra of automatically classified zooplankton from imaging systems: a case study with the ZooScan integrated system. *Methods in Oceanography*, **1–2**, 3–21.
- Vinogradov, M. (1962) Feeding of the deep-sea zooplankton. *Rapp. Pêches Cons. Perm. Int. Explor. Mer*, **153**, 114–120.
- Wiebe, P., Morton, A., Bradley, A., Backus, R., Craddock, J., Barber, V., Cowles, T. and Flierl, G. (1985) New development in the MOCNESS, an apparatus for sampling zooplankton and micronekton. *Mar. Biol.*, **87**, 313–323.
- Wishner, K. F., Seibel, B. A., Roman, C., Deutsch, C., Outram, D., Shaw, C., Birk, M. A., Mislan, K. et al. (2018) Ocean deoxygenation and zooplankton: very small oxygen differences matter. *Sci. Adv.*, **4**, eaau5180.
- Zhang, X. and Dam, H. G. (1997) Downward export of carbon by diel migrant mesozooplankton in the central equatorial Pacific. *Deep-Sea Res. II Top. Stud. Oceanogr.*, **44**, 2191–2202.