ZooProcess and Ecotaxa output from ZooSCANs of zooplankton collected along physical gradients during OAPS MOCNESS tows during R/V Oceanus northwest Atlantic 2011 cruise OC473 and R/V New Horizon northeast Pacific 2012 cruise NH1208 and imaged in 2021-2022

Website: https://www.bco-dmo.org/dataset/932252
Data Type: Cruise Results
Version: 1
Version Date: 2024-07-11

Project
» Quantifying the drivers of midwater zooplankton community structure (Zooplankton Gradients)

Contributors	Affiliation	Role
Blanco-Bercial, Leocadio	Bermuda Institute of Ocean Sciences (BIOS)	Principal Investigator
Maas, Amy	Bermuda Institute of Ocean Sciences (BIOS)	Co-Principal Investigator
Gossner, Hannah	Bermuda Institute of Ocean Sciences (BIOS)	Contact, Technician
York, Amber D.	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Abstract
This dataset consists of the imaging portion of the study described below and includes ZooProcess and Ecotaxa outputs from ZooSCANs performed of zooplankton collected during Multiple Opening-Closing Net and Environmental Sensing System (MOCNESS) tows during R/V Oceanus cruise OC473 in the Northwestern Atlantic in 2011 and R/V New Horizon cruise NH1208 in the Northeastern Pacific in 2012. It includes data for this project from Ecotaxa (export v1.0), an online machine-learning platform that assists in identifying organisms and particles. The dataset also includes particle measurements generated by ZooProcess software. Day and night stations were sampled between 0 to 1000m depths from 35 to 50 N in the northwest Atlantic in 2011, and from 35 and 50N along CLIVAR line P17N in 2012. These representative subsamples of the formalin-preserved zooplankton community from each net were imaged in 2021 and 2022. Project description: The objective of this study was to determine how environmental variables shape zooplankton community structure in the midwater. Our primary overarching hypothesis was that the abundance and size class distribution of the zooplankton community are decoupled and are influenced by different environmental variables. Furthermore, differences in zooplankton community composition and diversity in the observed distinct oceanic biogeographical provinces additionally influences both factors. Since zooplankton contributions to biogeochemistry are size dependent, standard descriptions of zooplankton community (biomass, which is a product of size and abundance) are insufficient to generate a predictive understanding of the role of zooplankton in biogeochemical cycles. The project uses particle imaging technology and metabarcoding of archived biological samples in conjunction with open access hydrographic data from two cruises conducted in the N. Atlantic and N. Pacific to test these hypotheses.

Table of Contents

- Coverage
- Dataset Description
  - Methods & Sampling
  - Data Processing Description
  - BCO-DMO Processing Description
  - Problem Description
- Data Files
- Supplemental Files
- Related Datasets
- Parameters
- Instruments
- Deployments
- Project Information

- Funding

Coverage

Location: N. Atlantic and N. Pacific
Spatial Extent: **N**:50.0913 **E**:-52.0707 **S**:33.5052 **W**:-134.979
Temporal Extent: 2011-08-13 - 2022-01-01

Dataset Description

Acknowledgment:

This dataset utilized samples and data funded by a previous NSF award (OCE-1041068) "Horizontal and Vertical Distribution of Thecosome Pteropods in Relation to Carbonate Chemistry in the Northwest Atlantic and Northeast Pacific (OAPS)" https://www.bco-dmo.org/project/2154.

Methods & Sampling

Location:

Northwestern Atlantic, 0 to 1000m depth, 35N, 52W to 50N, 42W along CLIVAR/WOCE line A20
Northeastern Pacific, 0 to 1000m depth, 50N, 150W to 35N, 135W along CLIVAR/WOCE line P17N

Methodology:

To obtain samples, a $1m^2$ Multiple Opening/Closing Net and Environmental Sensing System (MOCNESS; Wiebe et al., 1985) equipped with nine 150 m nets was deployed during the mid-day and mid-night on cruises carried out as described in https://www.bco-dmo.org/dataset/3546. Briefly, sampling was at consistent intervals including 1000-800, 800-600, 600-400, 400-200, 200-100, 100-50, 50-25, 25-0m at day/night stations from 35 to 50 N in the northwest Atlantic in 2011, and from 35 and 50N along CLIVAR line P17N in 2012.

Upon retrieval, the catch from each of the eight discrete nets were divided into splits. One-half of a sample was preserved in 95% ethanol, one quarter was preserved in 5% buffered formalin, and one quarter was used for live viewing and picking, and then preserved in 70% ethanol.

A representative subsample of the formalin-preserved zooplankton community from each net were imaged using a ZooSCAN ver. 4 at either 4,800 dpi with a narrow frame, or 2400 dpi with a large frame (following the methods in: Gorsky et al., 2010, Vandromme et al., 2012, as detailed in Maas et al. 2021). Scanning resolution was changed partway through the project in accordance with advice from Hydroptic (written communication). 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 ("d1"). The remainder of the sample was sieved through a 1-mm mesh sieve, and both size fractions ("d2" >1000um but excluding d1; "d3" between 150-1000um) were individually scanned. From these smaller size fractions, at least 1500 particles were scanned after subsampling using a Motoda splitter (Motoda, 1959), requiring generation of two separate scans for both size classes. This resulted in a total of five images per net.

ZooSCAN Image names:

Image names include: cruise#_mocnessID_net#_sizefraction_ and _a|b if a replicate and end in _1.tif

Multiple images of the same size fraction were sometimes taken to obtain a sufficient number of particles. These replicates are named a or b. If there is no replicate they don't have a letter in the image name. An a and b scan were always done for size classes d2 and d3. This was important because the split size is for the sum of a+b (e.g. if a is ¼ and b is ¼, the acq_sub_part will be 0.5).

Example of ZooSCAN image names:

ae1830_m13_n4_d3_a_1.tif [a replicate]
ae1830_m13_n4_d3_b_1.tif [a replicate]
ae1830_m13_n5_d1_1.tif [no replicate]

The "object_id" in this dataset (the particle identifier) is constructed the same way as the ZooSCAN image

name except it as an additional _# at the end. This additional number in the object_id is added by the ZooProcess software (Hydroptic, 2016).
e.g.
object_id: ae1614_m3_n1_d2_a_1_**100**
image_name: ae1614_m3_n1_d2_a_1.tif

Particle names:

Names for particles (objects) follow the pattern "CruiseID_MocnessID_NetNumber_ScanFraction" followed by "_1_XXX", with "1" being automatically added by the software to indicate no duplicates of that scan and "XXX" being the unique particle number within that scan.

Parameter (column name) nomenclature and data origin:
* (see the "Parameters" section which contains all column information for the ecotaxa output table)

Parameters (column names) beginning with "object" include basic identifying metadata input by the user as well as all particle measurement data generated by ZooProcess. Any parameters beginning with "object_annotation" parameters are added by Ecotaxa. Parameters that begin with "sample" are sampling metadata input by the user during the scanning process. "Process" parameters describe the software and assumptions or corrections input during the data processing. "Acq" describes the portion of the sample scanned (input by user) and provides some summary data about the scanned image.

Interpreting acq_sub_part

The acq_sub_part is the denominator of the fraction of the sample scanned (e.g. "8" would be 1/8 or 0.125). For this project, the acq sub part is the fraction of the whole size fraction. "A" and "B" scans of the same fraction are considered to be halves of the whole fraction and can be statistically lumped together (e.g. if d2_a and d2_b from the same net were each 0.25 ("4") of that size fraction, the acq sub part will be 0.5 ("2").

Instruments summaries:

MOCNESS

The Multiple Opening/Closing Net and Environmental Sensing System or MOCNESS is a family of net systems based on the Tucker Trawl principle. There are currently 8 different sizes of MOCNESS in existence which are designed for capture of different size ranges of zooplankton and micro-nekton. Each system is designated according to the size of the net mouth opening and in two cases, the number of nets it carries. The original MOCNESS (Wiebe et al, 1976) was a redesigned and improved version of a system described by Frost and McCrone (1974)(from MOCNESS manual). These nets allow for the discrete sampling of zooplankton at targeted depths while simultaneously capturing data on physical parameters. The MOCNESS used in all sampling for both cruises was a $1m^2$ (mouth size) rigged as below.

From OC473 Cruise Report (Lawson, et al. (2011), http://hdl.handle.net/1834/43091):

"The MOCNESS was equipped with eight 150-um mesh nets (nets 1-8; borrowed from URI) and one 333-um mesh net (net 0). The underwater unit used was #169. In addition to the standard temperature and conductivity probes the system also had a beta-type strobe-light unit for reducing avoidance of the nets by some zooplankton and possibly small fish. The strobe system has two units each with 12 LED sets (LUXEON Rebel LED) with peak output between 490-520 nm. Seven of the 24 LED sets were no longer working at the start of the sampling. The LEDs are powered by the MOCNESS battery and their pulse width, amplitude, flash rate period, and on/off are controlled by the MOCNESS software. For this cruise the pulse width was 2 ms, the relative amplitude was 99%, and the flash interval was 100 ms."

From NH1208 Cruise Report (Lawson, et al. (2012), http://hdl.handle.net/1834/43090):

"The MOCNESS was equipped with eight 150-um mesh nets (nets 1-8; borrowed from URI) and one 333-um mesh net (net 0). The system was equipped with the standard SeaBird temperature and conductivity probes (units #535 and #120 respectively). The underwater unit used was #169. In addition to the standard temperature and conductivity probes the system also had a beta-type strobe-light unit for reducing avoidance of the nets by some zooplankton and possibly small fish. The strobe system has two units each with 12 LED sets (LUXEON Rebel LED) with peak output between 490-520 nm. Seven of the 24 LED sets were no longer working at the start of the sampling. The LEDs are powered by the MOCNESS battery and their pulse width, amplitude, flash rate period, and on/off are controlled by the MOCNESS software. For this cruise the pulse width was 2 ms, the relative amplitude was 99%, and the flash interval was 100 ms. The strobe unit was only used for the first four tows, after which problems with blowing the underwater unit 5A fuse (a symptom typical

of strobe unit problems from the past) led us to disconnect it and not use it for the remainder of the tows."

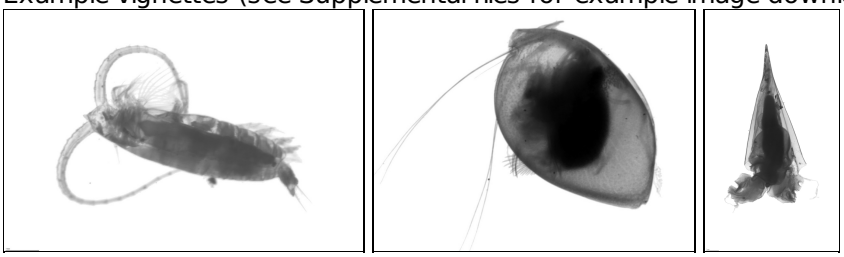
ZOOSCAN

The ZooSCAN (CNRS patent) system makes use of scanner technology with custom lighting and a watertight scanning chamber into which liquid zooplankton samples can be placed. The scanner recovers a high-resolution, digital image and the sample can be recovered without damage. These digital images can then be investigated by computer processing. While the resolution of the digitized zooplankton images is lower than the image obtained using a binocular microscope, this technique has proven to be more than adequate for large sample sets. Identification of species is done by automatic comparison of the image (vignette) of each individual animal in the scanned image with a library data set which may be built by the investigator for each individual survey or imported from a previous survey. The latest machine learning algorithm allows high recognition levels even if we recommend complementary manual sorting to achieve a high number of taxonomic groups. Scans for this dataset performed with a ZooSCAN (Hydroptic, HYDROPTIC_V4) running with Vuescan (version 9.5.24) and ZooProcess (version 8.22, ImageJ macro suite). Images were taken at either 4800dpi with a narrow frame or 2400dpi with a large frame.

Data Processing Description

Scans were processed using ZooProcess (version 8.22, ImageJ macro suite). The "Convert and process from RAW" function was used to separate particles into individual vignettes and generate a suite of measurements for each particle. "Doubles" (vignettes containing more than one particle) were manually separated in the software and reprocessed.

Example vignettes (see Supplemental files for example image download):



Processed scans and their corresponding metadata were then uploaded to Ecotaxa (Picheral et al, https://ecotaxa.obs-vlfr.fr/), where a training set was created using manually classified images from this project as well as existing validated images from other projects in the Sargasso Sea. Classification categories were chosen based on taxon of interest, identification level in previous projects, and known limitations of the software. Generally, broader level taxonomic groups are used. Identification of all particles was predicted, then manually validated.

Data versioning

The ecotaxa data included in this dataset version are "Ecotaxa Export version 1.0."

The ecotaxa export file will be updated in a versioned manner as validation of identification is completed. Ecotaxa output versions are as follows:

Version 1: No identifications, predicted or validated
Version 2: All identifications predicted
Version 3: All identifications validated

BCO-DMO Processing Description

Version 1:
* Data from source file ecotaxa_export_5421_20240614_1445_v1_0.tsv were imported into the BCO-DMO data system as the primary table for this dataset with "nan" values interpreted as missing data identifiers.
* After import, select columns designated to have the missing data identifier 99999 had a find and replace operation performed to turn 99999 values into system missing data identifiers. This was not done on import

since only some columns were described as using this identifier by the data providers. 99999 is within the valid numeric range for columns such as object_intden and object_area_exec so it is possible that if there were any real 99999 values, they were instead interpreted as missing data.

** In the BCO-DMO data system missing data identifiers are displayed according to the format of data you access. For example, in csv files it will be blank (null) values. In Matlab .mat files it will be NaN values. When viewing data online at BCO-DMO, the missing value will be shown as blank (null) values.

* There were two sets of lat lon columns called object_lat, object_lon. The metadata described these as the start and end lat,lon of the sampling event. So renamed -> object_lat_start, object_lon_start, object_lat_end, object_lon_end. The values were the same in both sets except some were null in the end set where the start set was 33.50517,-135.

* object_ISO_DateTime_UTC added from local date and time columns.

* extra trailing slash removed in object_link (all values were "http://www.zooscan.obs-vlfr.fr//"). This link may not be persistent for long-term curation, a citation for the site was added as a related publication along with the date it was accessed.

* The following blank columns were removed from this table to be consistent with columns in other ecotaxa table (BCO-DMO dataset 931883).
** object_annotation_status object_annotation_person_name, object_annotation_person_email, object_annotation_date, object_annotation_time, object_annotation_category, object_annotation_hierarchy, complement_info

* Parameters (column names) renamed to comply with BCO-DMO naming conventions. See https://www.bco-dmo.org/page/bco-dmo-data-processing-conventions (% in col name changed to "perc", trailing periods removed).

Problem Description

There is one missing net in this dataset- OC473_M14_N4 (200-400m depth bracket)- which was the result of a lost cod end during sampling. All other tows and nets were imaged normally. Some nets do not have a "d1" (>5000um) size fraction; this is because there were no organisms in this size fraction present.

Some chaetognaths and all pteropods were removed prior to imaging in association with the original OAPS and ancillary projects.

Depth ranges of nets were calculated by pulling the minimum and maximum depths of each net from the raw MOCNESS data, resulting in a small gap between net depth ranges due to sampling intervals in the raw data.

Considerable software issues were encountered while completing this project. Efforts were made to recover all relevant data and metadata, but there is a fair amount of metadata, particularly regarding the acquisition and processing steps, which is missing. All measurement and author-input metadata (especially context related) were checked and corrected as necessary.

Any data showing as a blank denotes an error in data generation or metadata retention in the zooprocess software pipeline. In future updates, this dataset will include annotation columns. Blanks in the annotation columns are due to no identification having been assigned yet.

[table of contents | back to top]

Data Files

File	
932252_v1_oaps_ecotaxa-and-zooprocess.csv	(Comma Separated Values (.csv), 644.44 MB) MD5:ade875dcc3042cf64779b21c61634e9e
Primary data file for dataset ID 932252, version 1	

[table of contents | back to top]

Supplemental Files

File
Example copepod vignette filename: copepod_6994.jpg (JPEG Image (.jpg), 289.71 KB) MD5:48a2abf039b09fd29dd60d2c8b5e6c41 Example vignette of one object. See methodology for more details about how these are used in Ecotaxa and how the vignettes are extracted from raw ZooScan images (see "Related Datasets" section for ZooScan images).
Example ostracod vignette filename: ostracod_female.jpg (JPEG Image (.jpg), 282.64 KB) MD5:d184d0302369c3e457dfdd328a46d5bc Example vignette of one object. See methodology for more details about how these are used in Ecotaxa and how the vignettes are extracted from raw ZooScan images (see "Related Datasets" section for ZooScan images).
Example pteropod vignette filename: thecosome1.jpg (JPEG Image (.jpg), 196.87 KB) MD5:30de45c826833e537998cc589e142ca7 Example vignette of one object. See methodology for more details about how these are used in Ecotaxa and how the vignettes are extracted from raw ZooScan images (see "Related Datasets" section for ZooScan images).

[table of contents | back to top]

Related Datasets

IsRelatedTo

Blanco-Bercial, L., Maas, A., Gossner, H. (2024) **ZooSCAN images of zooplankton collected along physical gradients during OAPS MOCNESS tows during R/V Oceanus cruise OC473 in the northwest Atlantic in 2011 and R/V New Horizon cruise NH1208 in the northeast Pacific in 2012 and imaged in 2021-2022.** Biological and Chemical Oceanography Data Management Office (BCO-DMO). (Version 1) Version Date 2024-04-15 doi:10.26008/1912/bco-dmo.865757.1 [view at BCO-DMO]
Relationship Description: ZooSCAN raw images that were analyzed to produce this dataset.

[table of contents | back to top]

Parameters

Parameter	Description	Units
object_id	Particle identification number. Typically cruiseID_mocnessID_net#_taxonomy_image#. [Source: User]	unitless
object_lat_start	Starting latitude of sample collection. [Source: User]	Decimal degrees
object_lon_start	Starting longitude of sample collection. [Source: User]	Decimal degrees
object_date	Date of sample collection. Local time zone (ADT/AST). [Source: User]	unitless

object_time	Start time of sample collection. Local time zone (ADT/AST). [Source: User]	unitless
object_ISO_DateTime_UTC	Datetime with timezone (start time of sample collection). Format ISO 8601, Time Zone UTC. [Source: User]	unitless
object_link	URL to Villefranche sur mer Quantitative Imaging Platform which hosts manuals summarizing image capture and upload. Autogenerated by software. [Source: Zooprocess]	unitless
object_depth_min	Minimum depth of sample collection (Zmin, shallowest net depth). [Source: User]	Meters
object_depth_max	Maximum depth of sample collection (Zmax, deepest net depth). [Source: User]	Meters
object_lat_end	Ending latitude of sample collection. [Source: User]	Decimal degrees
object_lon_end	Ending longitude of sample collection. [Source: User]	Decimal degrees
object_area	Surface area of the object. [Source: Zooprocess]	Square pixels
object_mean	Average grey value within the object. Pixel intensity value 0-255 (8 bit greyscale image).	pixel intensity
object_stddev	Standard deviation of the grey value used to generate the mean grey value. Pixel intensity value 0-255 (8 bit greyscale image).	pixel intensity
object_mode	Modal grey value within the object. Pixel intensity value 0-255 (8 bit greyscale image).	pixel intensity
object_min	Minimum grey value within the object (0=black). Pixel intensity value 0-255 (8 bit greyscale image).	pixel intensity
object_max	Maximum grey value within the object (255=white). Pixel intensity value 0-255 (8 bit greyscale image).	pixel intensity
object_x	X position of the center of gravity of the object within the vignette. [Source: Zooprocess]	Pixels
object_y	Y position of the center of gravity of the object within the vignette. [Source: Zooprocess]	Pixels

object_xm	X position of the center of gravity of the object's grey levels within the vignette. [Source: Zooprocess]	Pixels
object_ym	Y position of the center of gravity of the object's grey levels within the vignette. [Source: Zooprocess]	Pixels
object_perim	Length of the outside boundary (perimeter) of the object. [Source: Zooprocess]	Pixels
object_bx	X coordinate of the top left point of the smallest rectangle enclosing the object within the whole scan image. [Source: Zooprocess]	Pixels
object_by	Y coordinate of the top left point of the smallest rectangle enclosing the object within the whole scan image. [Source: Zooprocess]	Pixels
object_width	Width of the smallest rectangle enclosing the object. [Source: Zooprocess]	Pixels
object_height	Height of the smallest rectangle enclosing the object. [Source: Zooprocess]	Pixels
object_major	Primary axis of the best fitting ellipse for the object. [Source: Zooprocess]	Pixels
object_minor	Secondary axis of the best fitting ellipse for the object. [Source: Zooprocess]	Pixels
object_angle	Angle between the primary axis and a line parallel to the x-axis of the image. [Source: Zooprocess]	Decimal degrees
object_circ	Circularity. (4*pi*Area)/Perim^2. A value of 1 indicates a perfect circle, a value approaching 0 indicates an increasingly elongated polygon. . [Source: Zooprocess]	unitless
object_feret	Maximum ferret diameter, i.e. the longest distance between any two points along the object boundary. [Source: Zooprocess]	Pixels
object_intden	Integrated density. The sum of the grey values of the pixels in the object (i.e. = Area*Mean). . [Source: Zooprocess]	unitless
object_median	Median grey value within the object. Pixel intensity value 0-255 (8 bit greyscale image). [Source: Zooprocess]	pixel intensity

object_skew	Skewness (third order moment about the mean) of the histogram of the grey level values. [Source: Zooprocess]	unitless
object_kurt	Kurtosis (fourth order moment about the mean) of the histogram of grey level values. [Source: Zooprocess]	unitless
object_perc_area	Percentage of object's surface area that is comprised of holes, defined as the background grey level. [Source: Zooprocess]	Percentage
object_xstart	X coordinate of the top left point of the image within the scan. [Source: Zooprocess]	Pixels
object_ystart	Y coordinate of the top left point of the image within the scan. [Source: Zooprocess]	Pixels
object_area_exc	Surface area of the object excluding holes (=Area*(1-(%area/100)). [Source: Zooprocess]	Square pixels
object_fractal	Fractal dimension of the object boundary (Berube and Jebrak, 1999). [Source: Zooprocess]	unitless
object_skelarea	Surface area of skeleton in pixels. In a binary image, the skeleton is obtained by repeatedly removing pixels from the edges of objects until they are reduced to the width of a single pixel. (Berube and Jebrak, 1999). [Source: Zooprocess]	Pixels
object_slope	Slope of the grey level normalized cumulative histogram. [Source: Zooprocess]	unitless
object_histcum1	Grey level at 25% of the normalized cumulative histogram of grey levels. [Source: Zooprocess]	pixel intensity
object_histcum2	Grey level at 50% of the normalized cumulative histogram of grey levels. [Source: Zooprocess]	pixel intensity
object_histcum3	Grey level at 75% of the normalized cumulative histogram of grey levels. [Source: Zooprocess]	pixel intensity
object_xmg5	X position of the center of gravity of the object using a gamma value of 5. [Source: Zooprocess]	Pixels
object_ymg5	Y position of the center of gravity of the object using a gamma value of 5. [Source: Zooprocess]	Pixels

object_nb1	Number of remaining objects in the image after thresholding on level Histcum 1. [Source: Zooprocess]	unitless
object_nb2	Number of remaining objects in the image after thresholding on level Histcum 2. [Source: Zooprocess]	unitless
object_nb3	Number of remaining objects in the image after thresholding on level Histcum 3. [Source: Zooprocess]	unitless
object_compentropy	always set to 0, measurement no longer perfomed in Zooprocess. [Source: Zooprocess]	unitless
object_compmean	always set to 0, measurement no longer perfomed in Zooprocess. [Source: Zooprocess]	unitless
object_compslope	always set to 0, measurement no longer perfomed in Zooprocess. [Source: Zooprocess]	unitless
object_compm1	always set to 0, measurement no longer perfomed in Zooprocess. [Source: Zooprocess]	unitless
object_compm2	always set to 0, measurement no longer perfomed in Zooprocess. [Source: Zooprocess]	unitless
object_compm3	always set to 0, measurement no longer perfomed in Zooprocess. [Source: Zooprocess]	unitless
object_symetrieh	Bilateral horizonal symmetry index (Romagnan et al 2016). [Source: Zooprocess]	unitless
object_symetriev	Bilateral vertical symmetry index (Romagnan et al 2016). [Source: Zooprocess]	unitless
object_symetriehc	Symmetry of the object in relation to the horizontal axis after thresholding the grey level at Histcum 1 value. [Source: Zooprocess]	unitless
object_symetrievc	Symmetry of the object in relation to the vertical axis after thresholding at grey level Histcum 1 value. [Source: Zooprocess]	unitless
object_convperim	The perimeter of the smallest polygon within which all points of the object fit. [Source: Zooprocess]	Pixels

object_convarea	The area of the smallest polygon within which all points of the object fit. [Source: Zooprocess]	Pixels
object_fcons	Measure of contrast based on the texture feature descriptor (Amadasun and King, 1989). [Source: Zooprocess]	unitless
object_thickr	Thickness ratio; relationship between the maximum thickness of an object and the average thickness of the object excluding the maximum. . [Source: Zooprocess]	unitless
object_tag	Old variable which is no longer used (0 or 1; 1 if duplicate "tagged" object). [Source: Zooprocess]	unitless
object_esd	Equivalent spherical diameter. Sqrt(area/?). [Source: Zooprocess]	Pixels
object_elongation	Elongation index. Major axis/minor axis. [Source: Zooprocess]	unitless
object_range	Range of grey values (0-255). [Source: Zooprocess]	pixel intensity
object_meanpos	Relative position of the mean grey. (Mean grey- Max grey)/(Mean grey-Min grey). [Source: Zooprocess]	unitless
object_centroids	Difference between the mass and the grey-level object's centroids. Square root ((objGreyX - objX)^2 + (objGreyY - objY)^2). [Source: Zooprocess]	Pixels
object_cv	Coefficient of variation of grey values. 100*(stddev/mean). [Source: Zooprocess]	unitless
object_sr	Index of variation of grey values. 100*(stddev/max-min)). [Source: Zooprocess]	unitless
object_perimareaexc	Index of relative complexity of the perimeter. perimeter/area_exc. [Source: Zooprocess]	unitless
object_feretareaexc	Alternate elongation index. feret/area_exc. [Source: Zooprocess]	unitless
object_perimferet	Index of relative complexity of the perimeter. perim/feret. [Source: Zooprocess]	unitless
object_perimmajor	Alternate index of relative complexity of the perimeter. perim/major. [Source: Zooprocess]	unitless

object_circex	Circularity of the object excluding white pixels. (4*pi(?)*area_exc)/perimeter^2. [Source: Zooprocess]	unitless
object_cdexc	Distance between the mass and the grey-level object's centroids. (centroid^2)/area_exc. [Source: Zooprocess]	Pixels
sample_id	Name of sample (scan) from which the object originates (image_name without .tif extension). [Source: User]	unitless
sample_dataportal_descriptor	Optional descriptor. [Source: User]	unitless
sample_scan_operator	Initials of the individual who performed the scan. [Source: User]	unitless
sample_ship	Research vessel (if applicable) where sample was taken. [Source: User]	unitless
sample_program	Research program (if applicable) under which sample was taken. [Source: User]	unitless
sample_stationid	Station identifier (if applicable) at which sample was taken. [Source: User]	unitless
sample_bottomdepth	Recorded depth of the seafloor at the start of the tow. [Source: User]	Meters
sample_ctdrosettefilename	Associated CTD file name (if applicable). [Source: User]	unitless
sample_other_ref	Optional other reference notes. [Source: User]	unitless
sample_tow_nb	Number of tows, if more than one, combined to create sample. [Source: User]	Count
sample_tow_type	Type of tow profile used to collect sample. 1= oblique, 2 = horizontal, 3=vertical, 0= Other sampling method. [Source: User]	unitless
sample_net_type	Type of net used to collect the sample (e.g. MOCNESS, Reeve, Tucker, etc). [Source: User]	unitless

sample_net_mesh	Mesh size of the net used to collect the sample. [Source: User]	Microns
sample_net_surf	Area of net mouth. [Source: User]	Meters squared
sample_zmax	Maximum depth of tow or net for this sample. [Source: User]	Meters
sample_zmin	Minimum depth of tow or net for this sample. [Source: User]	Meters
sample_tot_vol	Total volume of water filtered while capturing this sample. [Source: User]	Meters cubed
sample_comment	Optional comment on this sample. [Source: User]	unitless
sample_tot_vol_qc	Quality flag for how volume of this tow or net was acquired. 1= RECORDED volume (flowmeter), 2= Calculated volume (using the mean volumne of other nets, 3= Estimated volume (net area * tow distance) . [Source: User]	unitless
sample_depth_qc	Quality flag for how the depth of this tow or net was acquired. 1= Measured by a depth sensor, 2= Calculated from cable length and angle, 3= Estimated from cable length. [Source: User]	unitless
sample_sample_qc	Quality flag for condition of the sample at time of scanning. Set of four digits defined below.; First digit is sample airtightness. 1=OK, 2=NOK (not ok); Second digit is sample richness. 1= Normal richness, 2=Very rich sample, 3=No plankton (almost) in sample; Third digit is sample condition. 1= Good condition , 2 = Dryed (no remaining liquid), 3= Rotton (loss of fixative); Fourth digit is disturbing elements. 1=No disturbing elements, 2=One of few large objects present in jar, 3=SOUP (phytoplankton, organic matter, clay/mud/mineral); . [Source: User]	unitless
sample_barcode	Barcode number assigned to sample if applicable. [Source: User]	unitless
sample_duration	Duration of the sampling tow. [Source: User]	Minutes
sample_ship_speed	Average ship speed during tow. [Source: User]	Knots
sample_cable_length	Maximum cable out during tow (m). [Source: User]	Meters
sample_cable_angle	Angle of the net cable during the tow (degrees from vertical). [Source: User]	Degrees

sample_cable_speed	Speed of the winch payout/retrival during the tow. [Source: User]	Meters per minute
sample_nb_jar	Number of the jar containing the sample. [Source: User]	unitless
sample_open	Undefined. [Source: User]	unitless
process_id	Process identifier. "zooprocess" and then the sample ID. [Source: Zooprocess]	unitless
process_date	Date the scan was performed. Local time zone (ADT/AST). [Source: Zooprocess]	unitless
process_time	Time the scan was performed. Local time zone (ADT/AST). [Source: Zooprocess].	unitless
process_img_software_version	Version of zooprocess software utilized. [Source: Zooprocess]	unitless
process_img_resolution	Resolution of the scan. [Source: Zooprocess]	Dpi
process_img_od_grey	Parameter not used- always nan. [Source: Zooprocess]	unitless
process_img_od_std	Parameter not used- always 0. [Source: Zooprocess]	unitless
process_img_background_img	Image name of background scan used. [Source: Zooprocess]	unitless
process_particle_version	Version of zooprocess software used to process particles. [Source: Zooprocess]	unitless
process_particle_threshold	8 bit coded grey level value to separate the objects from the background. [Source: Zooprocess]	pixel intensity
process_particle_pixel_size_mm	The number of pixels per millimeter with this dpi. [Source: Zooprocess]	Pixels per millimeter
process_particle_min_size_mm	Minimum size particle included in vignette extraction. [Source: Zooprocess]	Millimeters
process_particle_max_size_mm	Maximum size particle included in vignette extraction. [Source: Zooprocess]	Millimeters

process_particle_sep_mask	Whether or not to include a separation mask created during separation of touching vignettes. [Source: Zooprocess]	unitless
process_particle_bw_ratio	ratio of pixels from the image that are above the threshold value (i.e; objects of image noise). [Source: Zooprocess]	unitless
process_software	Software name, version, and date used to process this sample. [Source: Zooprocess]	unitless
acq_id	Name of the fraction or replicate being imaged (format fraction_cruise_mocness_net). [Source: Zooprocess]	unitless
acq_instrument	Type of instrument used to take original image (e.g. Zooscan, UVP, etc). [Source: User]	unitless
acq_min_mesh	The minimum mesh size/minimum hypothetical particle size of the fraction being imaged. [Source: User]	Microns
acq_max_mesh	The maximum mesh size/maximum hypothetical particle size of the fraction being imaged. [Source: User]	Microns
acq_sub_part	Portion of the sample in this image- denominator of the ratio (e.g. 1/64 will read "64"). [Source: User]	Integer
acq_sub_method	Method or splitter type used to subsample (e.g. Motoda, folsom). [Source: User]	unitless
acq_hardware	Hardware used to image the sample including version number (Zooscan model). [Source: Zooprocess]	unitless
acq_software	Software used to take the image- Vuescan version number. [Source: Zooprocess]	unitless
acq_author	Initials of individual who took the image. [Source: Zooprocess]	unitless
acq_imgtype	Type of image (e.g. zooscan, UVP). . [Source: Zooprocess]	unitless
acq_scan_date	Date that the scan was performed. Local time zone (ADT/AST). [Source: Zooprocess]	unitless
acq_scan_time	Time the scan was performed. Local time zone (ADT/AST). [Source: Zooprocess]	unitless

acq_quality	Parameter not used- always nan. [Source: Zooprocess]	unitless
acq_bitpixel	Flag- Vuescan coding of the grey resolution for the raw image. [Source: Zooprocess]	unitless
acq_greyfrom	Flag- The green channel of image from the scanner sensor is converted in grey by Vuescan for the image saving and processing. [Source: Zooprocess]	unitless
acq_scan_resolution	Flag- Vuescan coding of the resolution, utilized by Zooprocess to compute the pixel size and image resolution in dpi. [Source: Zooprocess]	unitless
acq_rotation	Flag- The raw image from the scanner is rotated before being saved. [Source: Zooprocess]	unitless
acq_miror	Flag- The raw image from the scanner is mirrored before being saved. [Source: Zooprocess]	unitless
acq_xsize	Horizontal size of the scan. [Source: Zooprocess]	Pixels
acq_ysize	Vertical size of the scan. [Source: Zooprocess]	Pixels
acq_xoffset	Scan frame X offset, from scanner factory calibration. [Source: Zooprocess]	Pixels
acq_yoffset	Scan frame Y offset, from scanner factory calibration. [Source: Zooprocess]	Pixels
acq_lut_color_balance	Indicates that the raw image normalisation will be done by Zooprocess. [Source: Zooprocess]	unitless
acq_lut_filter	If there is a filter in the acquisition LUT (all should be "no"). [Source: Zooprocess]	unitless
acq_lut_min	Minimum 16 bit coded grey level utilized for the 16 to 8 bit raw image conversion/normalisation by Zooprocess. [Source: Zooprocess]	pixel intensity
acq_lut_max	Maximum 16 bit coded grey level utilized for the 16 to 8 bit raw image conversion/normalisation by Zooprocess. [Source: Zooprocess]	pixel intensity

acq_lut_odrange	Optical Density range for the 16 to 8 bit raw image conversion/normalisation by Zooprocess (always 1.8). [Source: Zooprocess]	unitless
acq_lut_ratio	ratio applied to the acq_lut_max value for the 16 to 8 bit raw image conversion/normalisation by Zooprocess (always 1.15). [Source: Zooprocess]	unitless
acq_lut_16b_median	Measured median grey level of the raw image. [Source: Zooprocess]	pixel intensity

[table of contents | back to top]

Instruments

Dataset-specific Instrument Name	1m MOCNESS
Generic Instrument Name	MOCNESS1
Generic Instrument Description	The Multiple Opening/Closing Net and Environmental Sensing System or MOCNESS is a family of net systems based on the Tucker Trawl principle. The MOCNESS-1 carries nine 1-m2 nets usually of 335 micrometer mesh and is intended for use with the macrozooplankton. All nets are black to reduce contrast with the background. A motor/toggle release assembly is mounted on the top portion of the frame and stainless steel cables with swaged fittings are used to attach the net bar to the toggle release. A stepping motor in a pressure compensated case filled with oil turns the escapement crankshaft of the toggle release which sequentially releases the nets to an open then closed position on command from the surface. -- from the MOCNESS Operations Manual (1999 + 2003).

Dataset-specific Instrument Name	ZooSCAN ver. 4
Generic Instrument Name	ZooSCAN
Dataset-specific Description	http://www.hydroptic.com/index.php/public/Page/product_item/ZOOSCAN A representative subsample of the formalin-preserved zooplankton community from each net were imaged using a ZooSCAN ver. 4 at either 4,800 dpi with a narrow frame, or 2400 dpi with a large frame (following the methods in: Gorsky et al., 2010, Vandromme et al., 2012, as detailed in Maas et al. 2021). Scans for this dataset performed with a ZooScan (Hydroptic, HYDROPTIC_V4) running with Vuescan (version 9.5.24) and Zooprocess (version 8.22, ImageJ macro suite). Images were taken at either 4800dpi with a narrow frame or 2400dpi with a large frame.
Generic Instrument Description	Description excerpt from Hydroptic website http://www.hydroptic.com/index.php/public/Page/product_item/ZOOSCAN The ZooSCAN (CNRS patent) system makes use of scanner technology with custom lighting and a watertight scanning chamber into which liquid zooplankton samples can be placed. The scanner recovers a high-resolution, digitial image and the sample can be recovered without damage. These digital images can then be investigated by computer processing. While the resolution of the digitized zooplankton images is lower than the image obtained using a binocular microscope this technique has proved to be more than adequate for large sample sets. Identification of species is done by automatic comparison of the image (vignette) of each individual animal in the scanned image with a library data set which may be built by the investigator for each individual survey or imported from a previous survey. The latest machine learning algorithm allows high recognition levels even if we recommend complementary manual sorting to achieve a high number of taxonomic groups.

[table of contents | back to top]

Deployments

OC473

Website	https://www.bco-dmo.org/deployment/58720
Platform	R/V Oceanus
Report	http://hdl.handle.net/1834/43091
Start Date	2011-08-07
End Date	2011-09-01
Description	The primary objective of the proposed research is to quantify the distribution, abundance, species composition, shell condition, and vertical migratory behavior of oceanic thecosome pteropods in the northwest Atlantic and northeast Pacific, and correlate these quantities to hydrography and concurrent measurements of carbonate chemistry, including vertical and horizontal distributions of aragonite saturation. During OC473, the first cruise in the Atlantic, a combination of underway data collection and station activities will be conducted along a transect spanning 15 degrees of latitude (35° to 50° N) in the northwest Atlantic, employing six instrument packages: (1) a 1-m2 MOCNESS plankton net system; (2) a profiling Video Plankton Recorder / CTD package, including bottles for water sampling; (3) a deep (500m) towed broadband acoustic scattering system ; 94) a hull-mounted narrowband multi-frequency acoustic scattering system. It is possible that the hull mounted transducers will suffer from noise when the vessel is underway and so as a backup we will have a surface-towed sled with a backup complement of transducers; 5) an underway multi-parameter inorganic carbon analyzer and 6) a suite of chemistry-related instruments including a DIC auto-analyzer for discret bottle sample analysis, an alkalinity auto-titrator for bottle analysis and an Agilent spectrophotometer for discrete pH measurement. Supporting documentation: Cruise track image Cruise information and original data are available from the NSF R2R data catalog.

NH1208

Website	https://www.bco-dmo.org/deployment/58830
Platform	R/V New Horizon
Report	http://hdl.handle.net/1834/43090
Start Date	2012-08-09
End Date	2012-09-18
Description	The primary objective of this cruise was to quantify the distribution, abundance, species composition, shell condition, and vertical migratory behavior of oceanic thecosome pteropods in the northeast Pacific, and correlate these quantities to concurrent measurements of carbonate chemistry. Underway data collection and station activities were conducted on a transect running between 35 and 50N along CLIVAR line P17N. Six instrument types were used: (1) a 1-m2 MOCNESS plankton net system and a 1-m diameter Reeve net; (2) a profiling Video Plankton Recorder mounted on the CTD package that includes a Rosette system with Niskin bottles for water sampling; (3) a deep (500 meter) towed broadband acoustic scattering system; (4) a surface narrowband multi-frequency acoustic scattering system; (5) an underway multi-parameter inorganic carbon analyzer and a GO underway pCO2 system; and (6) a suite of chemistry-related lab instruments for bottle sample analysis including a DIC auto-analyzer, an alkalinity auto-titrator, and an Agilent spectrophotometer for pH measurement. The R/V New Horizon departed from Newport OR, and set a course for the transect start point at 50N 150W. Following instrument package test deployments over the continental shelf, the transect ran in a single zig-zag between the start point and the end at 35N 135W; a total of 34 stations were sampled along the transect, every 1/2 degree of latitude. In addition 10 other stations were sampled with a Reeve net for live experimental pteropods. The science party, divided into biology and chemistry teams conducted 24-hour operations. Cruise information and original data are available from the NSF R2R data catalog.

[table of contents | back to top]

Project Information

Quantifying the drivers of midwater zooplankton community structure (Zooplankton Gradients)

Website: https://www.nsf.gov/awardsearch/showAward?AWD_ID=1948162

Coverage: North Atlantic and North Pacific

NSF Award Abstract:

Processes in the midwater region below 200 m depth, also known as the twilight zone, represent a major unknown for the biology and chemistry of the ocean. Studies of animals drifting in the oceans, known as zooplankton, are scarce due to the difficulty and associated time and costs of sampling deep waters. The advent of automated image analysis and genetic tools is leading to a rapid increase in our knowledge of the diversity, abundances and size distribution of communities in shallow waters. However, our understanding of the deeper layers of the ocean is still in its infancy, and there are few studies that combine these three facets of the ecology of the zooplankton. The objective of this project is to leverage existing samples, obtained from previously NSF-funded research in the North Pacific and North Atlantic, to study how the abundances, diversity, and size distribution of zooplankton in the midwater vary with latitude and environmental factors. Automated image analyses provide information on abundance and size, and genetic analyses give unprecedented data on the diversity of the midwater community for the North Atlantic and the North Pacific, from subtropical to subarctic environments. This project provides high quality hands-on training opportunities for at least two undergraduate researchers and generates material for undergraduate and graduate courses. Two workshops train educators on the classroom use of the NSF-funded Biological and Chemical Oceanography Data Management Office (BCO-DMO) open access oceanographic data.

Recent advances in image analysis and metabarcoding of zooplankton communities via new data tools are an opportunity to generate quantitative and predictive relationships between environmental drivers and zooplankton diversity, abundances and size distribution. While this information is available for plankton in epipelagic regions, the focus of this study is on midwater communities, which remain poorly characterized. Obtaining these data is the first step towards a quantitative analysis that assesses the impact of the midwater community on biogeochemical cycles. The project uses archived samples from two cruises conducted in the N. Atlantic and N. Pacific to test hypotheses about how temperature, midwater hypoxia, primary productivity and biogeographic province shape the size class structure, biodiversity and behavior (diurnal vertical migration) of zooplankton communities. These newly-generated image and metabarcoding datasets of the mesozooplankton community from 0-1000 m are cross-comparable with other ocean regions. These data describe how migratory and midwater resident zooplankton communities are structured by environmental variables and demonstrate how this influences their biogeochemical contributions (specifically active flux and midwater attenuation of flux). Data tools generated for the image analysis in combination with metabarcoding has broad application in plankton ecology and allows metanalysis of other datasets. The project is complementary to ongoing national and international projects that seek to describe the function and structure of the midwater. In contrast to existing modeling and process projects, this project covers a moderately large geographic area and thus provides a strong comparative foundation for broader community-wide assessment of the function of zooplankton in the twilight zone.

This award reflects NSF's statutory mission and has been deemed worthy of support through evaluation using the Foundation's intellectual merit and broader impacts review criteria.

[table of contents | back to top]

Funding

Funding Source	Award
NSF Division of Ocean Sciences (NSF OCE)	OCE-1948162

[table of contents | back to top]