University of South Florida
Digital Commons @ University of South Florida

DIGITAL COMMONS
@ UNIVERSITY OF SOUTH FLORIDA

USF Tampa Graduate Theses and Dissertations
USF Graduate Theses and Dissertations

May 2024

# Abundance, Distribution, and Diversity of the Paralarval Cephalopod Community in the Northern Gulf of Mexico

Shannon Leah Riley
*University of South Florida*

Follow this and additional works at: https://digitalcommons.usf.edu/etd

 Part of the Other Oceanography and Atmospheric Sciences and Meteorology Commons

## Scholar Commons Citation

Riley, Shannon Leah, "Abundance, Distribution, and Diversity of the Paralarval Cephalopod Community in the Northern Gulf of Mexico" (2024). *USF Tampa Graduate Theses and Dissertations.*
https://digitalcommons.usf.edu/etd/10554

This Thesis is brought to you for free and open access by the USF Graduate Theses and Dissertations at Digital Commons @ University of South Florida. It has been accepted for inclusion in USF Tampa Graduate Theses and Dissertations by an authorized administrator of Digital Commons @ University of South Florida. For more information, please contact digitalcommons@usf.edu.

# Abundance, Distribution, and Diversity of the Paralarval Cephalopod Community in the Northern Gulf of Mexico

by

Shannon Leah Riley

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science
with a concentration in Biological Oceanography
College of Marine Science
University of South Florida

Co-Major Professor: Heather L. Judkins, Ph.D.
Co-Major Professor: Steven A. Murawski, Ph.D.
Glenn A. Zapfe, M.S.

Date of Approval:
April 25, 2024

Keywords: squid, octopus, assemblage, diel vertical migration, generalized additive model

Copyright © 2024, Shannon Leah Riley

**Acknowledgments**

This research was supported by the Anne and Werner Von Rosenstiel Fellowship in Marine Science, the Norman Blake Endowed Memorial Fellowship in Marine Science, and the Fish Florida Kaye Pearson Memorial Scholarship. I would like to thank all scientific and technical personnel who took part in the cruises during which my samples were collected as well as the staff of the Plankton Sorting and Identification Center at the Sea Fisheries Institute.

I would also like to thank my co-advisors, Dr. Heather Judkins and Dr. Steve Murawski, and committee member, Glenn Zapfe, for their support and guidance, as well as their flexibility and time as I wrap up my degree. I would like to thank Dr. Josh Kilborn for his excellent biometry and applied multivariate statistics courses, the latter of which was invaluable for a portion of this research. Thanks also go out to all current and former members of the Judkins and Murawski labs for their encouragement, feedback, and excellent company.

Finally, I would like to thank my friends and family for their love and support; and also commiserating with me as needed.

## Table of Contents

## List of Tables

## List of Figures

**Abstract**

The abundance, distribution, and diversity of cephalopod paralarvae from the upper water column (0-130 meters) in the northern Gulf of Mexico (GOM) was studied using samples collected during six ichthyoplankton surveys between 2009 and 2012. A total of 2240 cephalopod paralarvae belonging to 21 families were examined. Octopodidae, Enoploteuthidae, and Ommastrephidae were the most abundant taxa collected in the time series. Five taxa had significantly different abundances between some sampling periods, while no significant differences in abundance occurred between any GOM regions (northwest GOM shelf, north-central GOM shelf, West Florida Shelf, GOM slope, and GOM basin). The majority of taxa were found deeper in the water column at night than during the day. Loliginidae and Pyroteuthidae showed significant differences in abundances between day and night, while Octopodidae and Pyroteuthidae showed significant differences in abundances between depth bins at night, and no taxa showed significant differences in abundances between depth bins during the day. Some individual taxa also showed evidence of diel vertical migration, including Octopodidae and Pyroteuthidae. Statistical models showed that Enoploteuthidae abundance was affected by year, season, diel period, and depth bin, while Octopodidae abundance was affected by region, season, and fluorescence, and Loliginidae abundance was affected by light transmissivity, fluorescence, and year. Four clusters of taxa were identified, influenced mainly by the most abundant taxa. The composition and abundance of the paralarval community varied between years, seasons, diel periods, regions, and depths, although some levels of these factors (e.g., GOM slope and GOM basin) did not have significant differences. Abundance and distribution of paralarvae in epipelagic waters of the northern GOM is likely related to adult spawning timing

and location, vertical paralarval movement, hydrographic features, and circulation patterns dominated by the Loop Current.

# Chapter One: Abundance, Distribution, and Diversity of the Paralarval Cephalopod Community in the Northern Gulf of Mexico

## 1. Introduction

### 1.1. Cephalopods

Cephalopods are an important group of marine invertebrates consisting of octopods, squids, cuttlefish, and nautilus. Approximately 800 species of cephalopods exist today (Hoving et al. 2014). They range widely in size and are found in all areas of the ocean, from nearshore to deep-sea and from the tropics to polar regions. There are approximately 93 species of cephalopods representing 32 families in the Gulf of Mexico (GOM), though this number is likely to change with taxonomic revisions and discovery of new species (Judkins 2009). Some of the most commonly collected pelagic families are Enoploteuthidae, Cranchiidae, Pyroteuthidae, and Ommastrephidae (Judkins et al. 2017). The nearshore squid family Loliginidae is also abundant, as are several members of the octopus family Octopodidae (Judkins 2009).

Cephalopods are often very influential in marine food webs as both predators and prey (Piatkowski et al., 2001). Most species are generalist predators that consume a wide variety of prey, including zooplankton, fish, crustaceans, and other cephalopods (Anderson et al. 2008, Rodhouse 2013, Rosas-Luis et al. 2016, Guerra-Marrero et al. 2020, Portner et al. 2020). They are in turn predated upon by a range of marine megafauna, including fishes (e.g., commercially important tuna), marine mammals, and seabirds.

The position of cephalopods in marine food webs can be explored using food-web models, which often use stomach content analyses as inputs. These models show that the mean trophic level of squids is 3.65±0.39, but that their trophic level ranges between 2.35 and 4.42 (Coll et al. 2013). Stable isotope analysis can also be used to estimate trophic positions.

Using this method, a recent study calculated trophic positions between 3.1 and 4.7 for six GOM cephalopod species (Richards et al. 2023). The wide range of trophic positions, spanning two trophic levels, illustrates the range of roles cephalopods play within marine food webs. Cephalopods have similar trophic positions to zooplanktivorous fishes, but lower trophic positions than many micronektivore and piscivore fishes (Richards et al. 2023).

Additionally, species within the same family often show diverse trophic positions and diverse isotopic niche widths (Staudinger et al. 2019). Trophic position also increases with body size and ontogenetic shifts have been detected in some species (Staudinger et al. 2019). Squids in particular often have a large trophic impact on other parts of the food web and may be keystone species in some areas (Coll et al. 2013). Cephalopods are also important for transferring energy between trophic levels as they have high food conversion efficiency rates (Wells and Clarke, 1996).

Cephalopods are found throughout the water column from the surface to the deep-sea. Of the many cephalopod taxa that inhabit the mesopelagic zone, a large proportion travel up to the surface waters in some form of diel vertical migration (DVM). Cephalopods show a range of DVM patterns, including the most common, nyctoepipelagic synchronous vertical diel migration, in which animals spend daytime in the meso- and/or the bathypelagic zones and migrate to the epipelagic zone at night, and mesopelagic asynchronous vertical migration, in which animals are found primarily in the mesopelagic during the day and move into or between the epi- or the upper mesopelagic at night (Judkins and Vecchione 2020). DVM in cephalopods is likely driven by prey distribution and can be constrained by temperature requirements (Watanabe et al. 2006). DVM movements increase the contribution of cephalopods to carbon export from surface waters; thus they can be an important part of the biological carbon pump (Hidaka et al. 2001).

Although humans have consumed cephalopods for millennia in some regions, fishing has intensified in recent decades, illustrated by increased landings (Ospina-Alvarez et al. 2022). Worldwide cephalopod catches were 3.7 million metric tons in 2020 (FAO 2022), down from the

2014 peak of 4.9 million metric tons, but similar to the 2018 catches of 3.6 million metric tons (FAO 2020). This is only 4.1% of the global capture fisheries production, but cephalopods are considered one of the four most high-value groups (FAO 2022). The jumbo flying squid *Dosidicus gigas* had the highest single taxon catch, followed by other ommastrephid and loliginid squids, and cuttlefish (FAO, 2022).

In the United States, commercial harvest of cephalopods centers around three species: California market squid (*Doryteuthis opalescens*), longfin squid (*Doryteuthis pealeii*), and shortfin squid (*Illex illecebrosus*) (NOAA Fisheries 2020). These fisheries are managed under state or federal management plans which dictate catch quotas. Combined landings for the three species were 62,358 metric tons in 2020 (NOAA Fisheries 2022). Other species of cephalopods, including octopus, the Atlantic brief squid (*Lolliguncula brevis*), and the arrow squid (*Doryteuthis plei*), are also caught commercially, but in much lower quantities (Voss and Brakoniecki 1985, NOAA Fisheries 2022). The United States does not have any commercial cephalopod fishery within the GOM, but Mexico has a commercial fishery for octopus (Luis et al. 2020). Additionally, recreational and/or artisanal fishing for cephalopods occurs throughout the GOM.

Many of the commercially important cephalopods discussed above are found in the neritic zone, especially those that have been fished the longest. Early research on cephalopods was almost entirely focused on these common neritic species, especially the loliginids and some octopods, as they were the easiest to access (Arkhipkin 2004). Research on pelagic cephalopods has increased in recent years, notably for the commercially valuable ommastrephids. However, even general biological knowledge is still lacking for many species, especially for deep-sea cephalopods that spend most of their lives in the midwater column, such as the genera *Cirroctopus* and *Magnapinna*, due in large part to the inaccessibility of their environment (Hoving et al. 2014). In the GOM, nearshore species are well-studied, while research on pelagic species is slowly increasing (Voss 1954, Judkins 2009, Sutton et al. 2022).

## 1.2. Paralarvae

The first stage after hatching in the life cycle of most cephalopods is the paralarval stage, during which the animals are planktonic (Young and Harman 1988). Cephalopods that do not live in the water column at all during their lives, such as *Octopus maya* and *Sepia officinalis,* are considered to lack a paralarval stage (Vidal and Shea 2023). Termed “paralarvae” instead of “larvae” because they are very similar to the adult form and do not undergo metamorphosis, these young cephalopods have little locomotory control over their horizontal spatial position, similar to other members of the zooplankton community (Young and Harman 1988, Power 1989, Bartol et al. 2008, Vidal and Shea 2023). While paralarvae show most structural features of adults, they do have morphological specializations for a planktonic lifestyle that distinguish them from older conspecifics, including large chromatophores, rudimentary paddle-shaped fins, and few or no photophores (Vidal and Shea 2023). The duration of the paralarval stage varies widely between different taxa, from days to months [e.g., *Sepiola atlantica* (Sepiolidae): six days (Jones and Richardson 2010) and *Robsonella fontaniana* (Octopodidae): 70 days (Uriarte et al. 2010)]. Due to their rapid growth, short paralarval phase, and patchy distribution, paralarvae are relatively rare in the plankton community, which makes them challenging to study.

Research focusing on paralarvae, particularly in the wild, has expanded considerably in the last two decades. A number of studies on the distribution and abundance of paralarvae have been conducted in and around the Gulf of California, off the coast of the Iberian Peninsula, and off the coast of Brazil (e.g., Haimovici et al., 2002; Moreno et al., 2009; Roura et al., 2016; Ruvalcaba-Aroche et al., 2018; Araújo and Gasalla, 2019; Martínez-Soler et al., 2021). However, many regions are still lacking in knowledge about the paralarval community and there is a dearth of information about most taxa at the paralarval stage, especially for species that are not fished commercially.

As zooplankton, paralarvae are often distributed higher in the water column than adults. Myopsid squid paralarvae are found in neritic waters, as are the paralarvae of many demersal

octopods (Moreno et al. 2009). Oegospid squid paralarvae are often found further offshore, along with the paralarvae of pelagic octopods. Many of these are distributed in the first few hundred meters of the water column (e.g., Bower and Takagi, 2004; Castillo-Estrada et al., 2020).

Some species of cephalopod undergo DVM as paralarvae, although vertical migration is less common in paralarvae than adults. Some species vertically migrate at one life stage, but not at another. For example, Bower and Takagi (2004) found that *Gonatopsis borealis* and *Gonatus spp*. paralarvae did not show any evidence of diel vertical migration, but Watanabe *et al.* (2006) found that adults of these taxa underwent diel vertical migration. In contrast, both *Abralia redfieldi* paralarvae and adults showed clear evidence of DVM (Castillo-Estrada et al. 2020, Judkins and Vecchione 2020). However, Castillo-Estrada et al. (2020) found evidence of DVM in *A. redfieldi* in only one of two years of study, suggesting that more research is needed to confirm any migratory behavior.

Paralarvae are difficult to study because they tend to be patchily distributed, so a greater number of samples is required to obtain adequate sample sizes for robust analysis of abundance and distribution. Net type and collection depth are also important. Many species of cephalopods are found in near-surface waters as paralarvae, requiring use of a net capable of surveying this habitat. The main gear used to collect paralarval cephalopods is the bongo net, which is used for depth-integrated oblique hauls (e.g., Moreno et al., 2009). Sampling with a focus on vertical distribution frequently uses a Multiple Opening/Closing Net and Environmental Sensing System (MOCNESS) (e.g., Castillo-Estrada et al., 2020). On occasion, a manta net is the optimal choice, as it is best for surveying species frequently found at or near the surface, such as the *Sthenoteuthis oualaniensis - Dosidicus gigas* complex (e.g., Vecchione, 1999).

Cephalopod paralarvae are challenging to identify morphologically at the smallest sizes; many species are between one and two mm at hatching (e.g., *Abralia trigonura*: 1 mm, *Ommastrephes bartramii*: 1.3 mm, *Octopus insularis*: 1.7 mm) (Young and Harman 1985,

Villanueva et al. 2016). At these body sizes morphological characteristics such as funnel-locking mechanisms and tentacle clubs are so small that it can be difficult to pick out features using a standard dissecting microscope. Because of this, some studies make use of other types of microscopy, such as scanning electron microscopy and autofluorescence microscopy (Ramos-Castillejos et al. 2010, Metz et al. 2015, Fernández-Álvarez et al. 2018). Additionally, many useful features for distinguishing between species in the same family or genus do not appear until later in the animal's development, for example, within Ommastrephidae, the species *Sthenoteuthis oualaniensis* and *Dosidicus gigas* in the eastern Pacific cannot be distinguished by morphological identification until they attain mantle lengths of about 4 mm, at which point *S. oualaniensis* develops ocular and intestinal photophore while *D. gigas* does not (Ramos-Castillejos et al. 2010).

These challenges are common with even the most intact specimens. For specimens that have been damaged during capture, identification becomes even more difficult, especially if a useful body part, such as the tentacle club, has been separated from the body. Gelatinous species are often torn apart, and even more muscular species may lose arms, tentacles, or eyes or have the mantle everted. Due to these obstacles, taxa with especially distinctive features, such as the proboscis of paralarval ommastrephids, are often positively identified more frequently than those without such features. This suggests potential biases in the results of studies based on morphological identification. Taxa with distinctive features, such as the ommastrephids mentioned above, may be reported as more abundant, but this could be simply because they are easy to identify, while other taxa are left as "unidentified." The challenge of identifying paralarvae morphologically means that many studies are only the family level (e.g., Haimovici et al., 2002; Ruvalcaba-Aroche et al., 2018; Araújo and Gasalla, 2019). DNA barcoding can be a useful alternative to morphological identification as it often allows for species-level identification of even damaged specimens (Taite et al. 2020).

In the offshore waters of the GOM, one can expect to find oegopsid squid and pelagic octopod paralarvae. At nearshore locations, one can expect to find myopsid squid paralarvae and the paralarvae of nearshore octopuses. The presence of paralarvae is dependent on the spawning period of the adults; however currently little is known about spawning patterns of cephalopods specifically within the GOM, especially for pelagic species. Although many studies have been conducted on sub-adult and adult cephalopods in the GOM, to date, very few have focused on offshore (i.e., non-loliginid) paralarvae. Sluis et al. (2021) studied the distribution and abundance of squid paralarvae in the northern GOM. Goldman and McGowan (1991) studied the distribution and abundance of ommastrephid squid paralarvae in the Florida Keys, in the far eastern portion of the GOM. Santana-Cisneros *et al.* (2021a; 2021b) studied octopod paralarval identification and dispersion in the southern GOM. Guarneros-Narváez *et al.* (2022) studied the distribution, abundance, and genetic structure of loliginid squid paralarvae in the southeastern GOM. García-Cordova et al. (2023) examined the distribution of myopsid and oegopsid squid paralarvae in the southeastern GOM.

### 1.3. Gulf of Mexico

The GOM is a semi-enclosed basin considered a marginal sea of the Atlantic Ocean. It is connected to the Caribbean Sea via the Yucatán Strait and to the western Atlantic Ocean through the Straits of Florida. Circulation in the GOM is driven by the Loop Current, part of the Gulf Stream system (Candela et al. 2002, Kourafalou et al. 2017). The Loop current flows into the GOM through the Yucatan Channel as the Yucatan Current (Candela et al. 2002), then flows clockwise around the eastern GOM, exiting through the Straits of Florida, where it becomes the Florida Current (Kourafalou et al. 2017). The northern extent of the Loop Current varies throughout the year, reaching between 24°N and 29°N (Maul 1977, Leben 2005). At its farthest northern extension, the current almost reaches the Mississippi Delta (Maul 1977, Schiller and

Kourafalou 2014). The longitudinal extent also varies; the Loop Current enters the GOM at around 86°W but the westward extension of the anticyclonic flow can reach 93°W (Leben 2005).

At irregular intervals (between two weeks and 18 months), eddies split off from the Loop Current and travel to the west (Sturges and Leben 2000, Leben 2005). These Loop Current Eddies (LCEs) are large (average area: 38,768 $km^2$) warm-core mesoscale anticyclones (Leben 2005). LCEs are important in the westward transport of heat and salt, and the redistribution of river plumes (Elliott 1982, Morey et al. 2003). Surface chlorophyll concentration is lower in LCEs than in common GOM water; however, primary production in LCEs is higher than in the surrounding waters (Damien et al., 2021). Eddies can also transport zooplankton, sometimes large distances, due to long retention times (Condie and Condie 2016). Differences in plankton retention occur with depth, suggesting that plankton communities within the eddy may vary depending on its depth structure (Condie and Condie 2016). Warm-core eddies in the northwestern Atlantic Ocean support diverse species assemblages of paralarvae (Taite et al. 2020).

The Mississippi River, one of the largest rivers in the world by amount of water discharged ($15^{th}$ largest, National Park Service, 2022), has a large impact on the GOM. It is an important source of freshwater and terrigenous sediments and nutrients. Variability in the extension of the river plume is dependent on river discharge and alongshore winds (da Silva and Castelao 2018). Maximum offshore extension is typically during summer, with minimum extension in fall/winter (da Silva and Castelao 2018). The continental shelf is always under the influence of the river plume, but during summer the plume can extend beyond the shelf break (approximately 200 to 250 kilometers) and in anomalous years can even reach 350 kilometers from the coast due to interactions with the Loop Current and LCEs (da Silva and Castelao 2018).

The Mississippi River plume has a large influence on biomass at various trophic levels of the pelagic food web. Lohrenz *et al.* (1990) found that primary production was enhanced at the

interface of the plume and Gulf Common Water, due in part to the combination of light availability, nutrient concentration, and salinity. Microzooplankton grazing rates were also highest at intermediate stations between the plume and more standard GOM waters (Liu and Dagg 2003). Overall zooplankton biomass, and specifically copepod abundance, has been found to be highest in the Mississippi Outflow region (Ortner et al. 1989).

The GOM experiences seasonality in several environmental conditions, including temperature. The surface temperatures are uniformly warm in the summer (~29°C and higher) and cooler in the winter, with the coldest surface temperatures along the northern edge (~19°C), while temperatures remain higher further south (up to ~26°C) (Allard et al. 2016). Spring and fall temperatures are transitional, with spring temperatures slightly higher than winter and fall temperatures slightly lower than summer (Allard et al. 2016). Offshore sea surface salinity is fairly constant throughout the year, with lower salinity waters present in fall and winter in nearshore areas (Allee et al. 2012). The near-shore region of the northern GOM experiences seasonal hypoxia in bottom-waters, but near-surface oxygen concentrations are relatively constant throughout the GOM over the seasonal cycle (Garcia 2014). In general, chlorophyll concentration is highly variable throughout the GOM (Pasqueron De Fommervault et al. 2017). Chlorophyll concentrations are highest in coastal waters where riverine input increases nutrient levels. In terms of the seasonal cycle, chlorophyll concentrations are highest in the winter and lowest in the summer (Pasqueron De Fommervault et al. 2017). Spring and fall have average chlorophyll concentrations closer to the summer values (Pasqueron De Fommervault et al. 2017).

### 1.4. Objectives

This study will explore multiple aspects of cephalopod paralarvae ecology to better understand the abundance and distribution in the northern GOM and along the West Florida Shelf. Four main objectives will be addressed by the present study: 1) compare abundance and

spatial distribution of paralarvae across different time periods and regions, 2) compare day and night vertical abundance distribution of paralarvae in the upper water column, 3) investigate the influence of environmental, spatial, and temporal factors on paralarval abundance, and 4) quantify the diversity of the paralarval community to determine spatial and temporal differences.

Paralarval abundance in the GOM is likely to vary spatially and temporally, which is mainly dependent on spawning times and locations, as well as water currents. Additionally, paralarvae are likely affected by environmental conditions including water temperature and salinity. As zooplankton, the abundance of paralarvae may also vary with chlorophyll concentration and sample collection depth. Paralarval abundance is also affected by predation as they are consumed by a variety of predators (York and Bartol 2016). Examining these factors will allow inferences to be made about the conditions they experience and potential exposure to anthropogenic activities. The results of this research will also fill a regional gap in knowledge of paralarval cephalopod ecology by considering a wide spatial area and multiple taxa. Additionally, these results are a baseline for future studies to compare the paralarval communities in different locations or times as well as identify possible changes in the northern GOM paralarval community.

## 2. Methods

### 2.1. Sample Collection

Cephalopod paralarvae were collected by the Southeast Area Monitoring and Assessment Program (SEAMAP) during their annual plankton surveys in the northern GOM. SEAMAP is a joint state/federal/university program with partners including the NOAA Southeast Fisheries Science Center and the Florida Wildlife Research Institute that has been conducting surveys in the GOM since 1981. The samples used in this study were collected on surveys conducted in winter (January-March), spring (April-May), and early fall (August-September). Different areas were targeted in different seasons depending on SEAMAP objectives: mainly the

shelf break in the winter, with some sampling over the inshore continental shelf and offshore waters, offshore waters in spring, and continental shelf waters in fall. Individual sites were sampled in one, two, or all seasons. Stations were sampled regardless of time of day, so plankton collections were made around the clock, resulting in both daytime and nighttime samples (Southeast Area Monitoring and Assessment Program 2016). A CTD cast was conducted at the end of every station to obtain temperature, salinity, dissolved oxygen, transmissivity, and fluorescence profiles. The continuously sampled CTD data were binned to produce one observation per meter of deployment.

The zooplankton sampling was carried out using a MOCNESS with a 1 $m^2$ opening and nine 0.505 mm mesh nets. The first net (net 0) is open while the sampling gear descends through the water column and is generally discarded from analysis as it is an aggregate sample of the entire sampled portion of the water column. The remaining eight nets are opened and closed at specific, sequential depths to collect plankton from discrete depth bins. Depth bins were determined by water column depth; at shallow stations (<65 m), 10 m depth bins were used, while at stations with water depth greater than 65 m, 20 m depth bins were used. Each net is considered an individual sample. Therefore, each survey site has multiple samples collected on the same day. The MOCNESS records the opening and closing depths of each net and the MOCNESS software calculates the volume filtered by each net.

Samples were initially preserved in 95% ethanol after collection and transferred to fresh ethanol after 24 hours. After the cruises, samples were sorted and cephalopods were removed, counted, and placed in separate vials at the Sea Fisheries Institute, Plankton Sorting and Identification Center (ZSIOP), in Szczecin, Poland. Each vial contained the cephalopods collected in a single net.

This study examined samples collected between 2009 and 2012 during six surveys (Table 1). Samples were conducted in winter (late January through mid-March), spring (May), and early fall (late August through September). Eighty-nine individual sites were sampled during

the surveys (Figure 1). The number of sites occupied per survey ranged from 9 (early fall 2009) to 45 (winter 2009) (Table 1). The number of nets used per site in the SEAMAP sampling ranged from three to eight, resulting in between 60 and 246 samples collected per survey (Table 1). At deeper stations, sampling was conducted from 130 m depth to the surface, while at shallower stations, sampling was from near bottom to the surface. The volume of water filtered by the nets was typically between 200 and 400 $m^3$, where larger volumes filtered generally indicate a larger depth range sampled.

Table 1. Number of sites occupied and number of samples (not including net zeroes) collected in each survey used in this study.

| | *Winter 2009* | *Early Fall 2009* | *Early Fall 2011* | *Winter 2012* | *Spring 2012* | *Early Fall 2012* |
|---|---|---|---|---|---|---|
| *Sites Occupied* | 45 | 9 | 12 | 17 | 23 | 11 |
| *Samples Collected* | 246 | 61 | 60 | 94 | 115 | 61 |

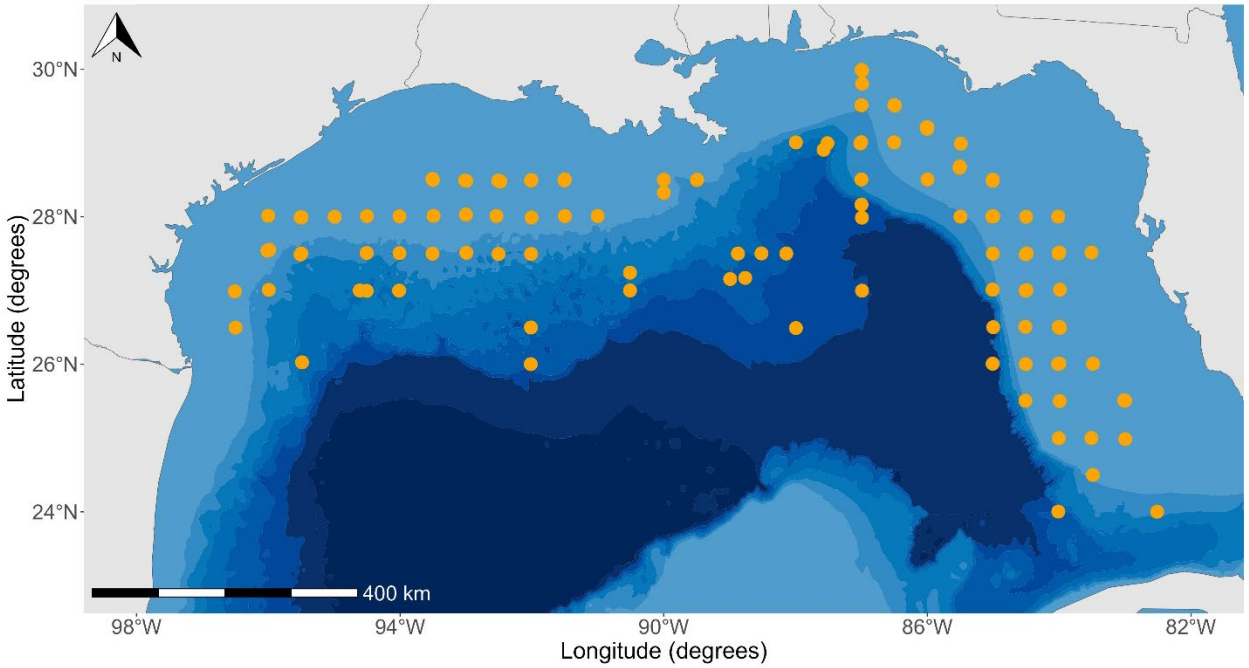


Figure 1. Sampling sites in the GOM (2009-2012, 6 cruises).

## 2.2. Specimen Identification

Specimens were identified to the lowest taxonomic level possible based on identification guides by Sweeney et al. (1992), Vecchione et al. (2001), and Vidal, Shea, and Judkins (unpubld). Damaged or small (generally less than two mm mantle length) specimens often could not be identified and were classified as “unidentified squid” or “unidentified cephalopod.” Paralarvae were measured for dorsal mantle length to the nearest 0.1 mm using digital calipers and weighed to the nearest 0.01 g. Paralarvae were split into groups based on identification and archived in vials with one vial per taxon per sample. If a large number of paralarvae from a single taxon were present in one sample, they were split into multiple vials for archiving.

## 2.3. Data Treatment

For each sample, the total depth covered by the net in meters was calculated, the depth bin sampled was identified, and the median point of the depth bin was calculated. Each sample was assigned to one of five regions based on Wilkinson et al. (2009): Northwestern GOM shelf, North-central GOM shelf, West Florida Shelf, GOM slope, and GOM Basin (Figure 2). These regions were selected from the three GOM regions (level I), then combining the seafloor geomorphological regions (level II) based on distance from coast (shelf, slope, other), and then splitting the shelf region into three regions based on coastal regions (level III) (Wilkinson et al. 2009).

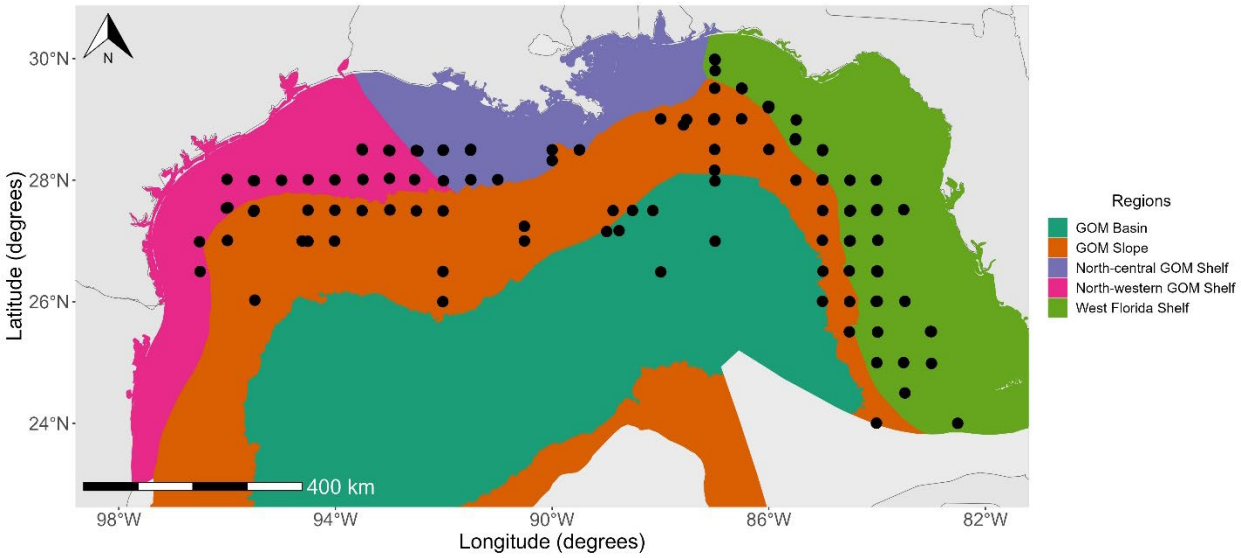


Figure 2. Sampling sites in the five GOM regions considered in this study.

2.4. Analysis

All analyses were carried out using R version 4.3.1 and RStudio (Posit team 2023, R Core Team 2023). Abundance modeling was performed using the R packages ape (v5.7-1; Paradis and Schliep 2019), car (v3.1-2; Fox and Weisberg 2018), mgcv (v1.8-42; Wood 2017), and mgcViz (Fasiolo et al. 2020). Community structure analyses and visualizations were carried out using the R packages BiodiversityR (v2.15-4; Kindt and Coe 2005), caret (v6.0-94; Kuhn 2008), ggordiplots (v0.4.3; Quensen et al. 2024), labdsv (v2.1-0; Roberts 2023), RVAideMemoire (v0.9-83-3; Herve 2023), and vegan (v2.6-4; Oksanen et al. 2022). A full list of packages used in this study is provided in Appendix 1.

Samples with extremely high counts of individual taxa were not removed from any analyses as they are real data points and their effect on the mean abundance is reflected in the standard deviation/standard error. All taxa were included in the analyses except for the abundance modeling, which included only three taxa with high abundance and frequency of occurrence. Unidentified individuals were not included in analyses, with one exception: all

paralarvae, including unidentified specimens, were combined into a single category (Cephalopoda) to include as a comparative group in the analyses.

*2.4.1. Abundance and Distribution*

The abundance of each taxon was calculated for all sampling periods and regions. The periods were: all cruises, 2009, 2011, 2012, winter 2009, early fall 2009, winter 2012, spring 2012, and early fall 2012. The regions were the Northwestern GOM shelf, North-central GOM shelf, West Florida Shelf, GOM slope, and GOM basin (Figure 2). The abundance of paralarvae in each sample was calculated by dividing the count of paralarvae (all paralarvae together and individual taxa) by the volume of water filtered by the net and multiplying by 1000. The general prevalence of each taxon as well as all paralarvae together was then quantified using the following five metrics: 1) total number of individuals of the specified taxon collected in a given sampling period or region, 2) mean abundance (number of paralarvae per 1000 $m^3$) and standard deviation of the specified taxon in a given period or region, 3) the percentage of all individuals (including unidentified) collected in a given sampling period or region belonging to the specified taxon, 4) the frequency of occurrence, or percent of samples collected in a given sampling period or region in which the specified taxon occurred, and 5) the number of samples collected in a given sampling period or region in which the specified taxon occurred.

Mean abundance was compared between different sampling periods and sampling regions for all taxa. The Wilcoxon test (used when two levels of grouping factor were present) or Kruskal-Wallis test (used if more than two levels of grouping factor were present) was used to test for significant differences in mean abundance between years, seasons, seasons in 2009, seasons in 2012, and regions. If the Kruskal-Wallis test was used and produced a significant p-value, Dunn's test with Holm's p-value adjustment was used for *post hoc* multiple comparison testing on all possible pairs to determine which pairs of the grouping factor had significantly

different mean abundances. Frequency of occurrence and mean abundance were considered in conjunction to identify the most common/abundant taxa.

*2.4.2. Vertical Distribution*

The number of samples collected in each depth bin was calculated (Figure 3) and the five most frequently sampled bins along with the most frequently sampled bin deeper than 100 m were selected for the vertical distribution analysis. Due to variation in collection depths, each sample was assigned to one of these six depth bins if its opening and closing depths fit within a depth bin. Samples that did not fit in any of the bins were removed from this portion of the analysis. Although the samples were collected at different times and in different locations, they were all used to assess the vertical distribution of the paralarvae. All metrics were calculated for each taxon individually as well as all specimens combined into one group. Before analyzing the mean abundances, the spread of abundances was visualized for each taxon.

Each taxon was then tested to identify taxa with significant diel differences in near-surface (0-130 m) abundance. All samples containing at least one individual of the specified taxon were subset from the full set of samples and the abundance of the taxon in each sample was calculated as in the previous section. Then the samples were split between day and night and the average abundance and standard deviation at each time of day were calculated. The Wilcoxon test was used to determine if there was a significant difference between the average abundance during the day versus at night ($\alpha = 0.05$). Samples from different depth bins were all considered together here.

This process was repeated to test for significant differences in abundance among depth bins. In this case, day and night samples were analyzed together. The Kruskal-Wallis test was used for taxa that were collected in more than two depth bins, while the Wilcoxon test was used for taxa that were collected in only two depth bins. If a taxon had a significant difference in mean abundance between depths ($\alpha = 0.05$) from the Kruskal-Wallis test, the Dunn test with Holm's p-

value adjustment was used for follow-up pairwise testing to identify pairs of depth bins with significantly different mean abundances.

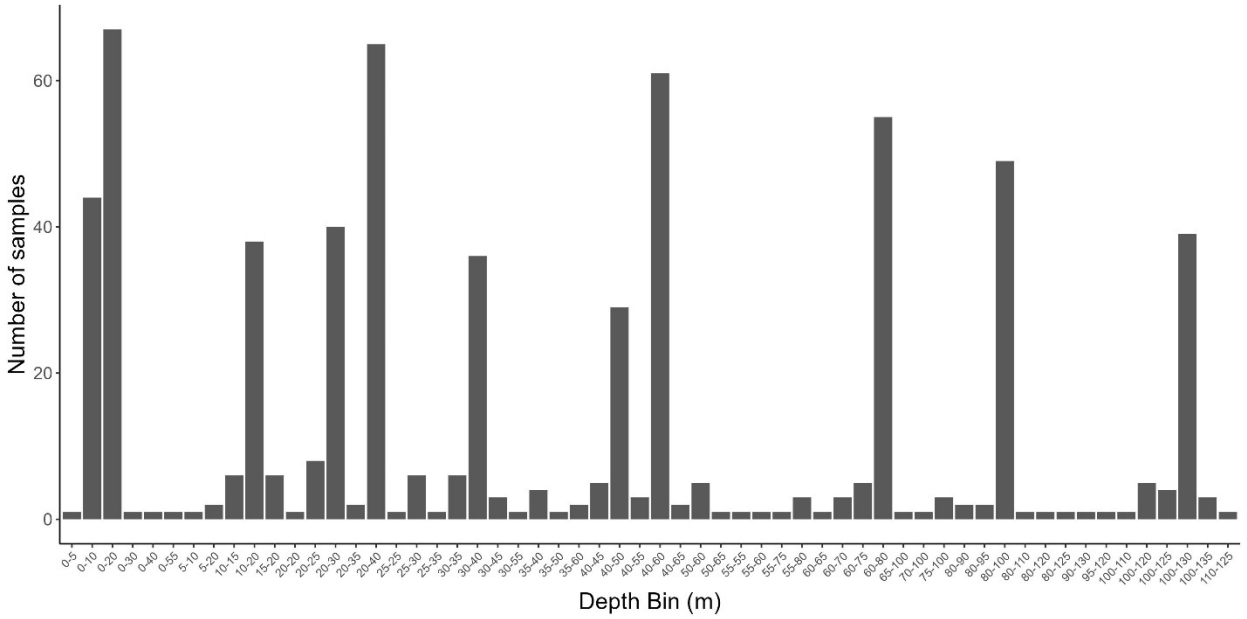


Figure 3. Number of samples collected in each depth bin.

Mean abundance was calculated separately for daytime and nighttime samples for each taxon in each depth bin. Significant differences among depth bins were identified using the Kruskal-Wallis or Wilcoxon test, and the Dunn test with Holm's p-value adjustment was used for follow-up testing.

The daytime and nighttime abundance-weighted mean depths were calculated for each taxon following Muhling et al. (2013). This metric identifies the center of distribution for each taxon. The following equation was used to calculate the abundance-weighted mean depth for each time of day separately for each taxon:

$$Y = \frac{\sum_{i=1}^{n} w_i X_i}{\sum w_i}$$

where $Y$ is the abundance-weighted mean depth at a specified time of day for a specified taxon, $X$ is the median depth of the sample's depth bin, and $w_i$ is the abundance of sample $i$. For each taxon, the day and night abundance-weighted mean depths were compared to make inferences about possible diel movement.

*2.4.3. Abundance Modeling*

Generalized additive modeling was used to examine the relationship between paralarval abundance (as counts) and a range of environmental, spatial, and temporal variables. Three common (high abundance and high frequency of occurrence) taxa were selected for abundance modeling: one oegopsid family (Enoploteuthidae), one myopsid family (Loliginidae), and one octopod family (Octopodidae). Each taxon was modeled individually. The explanatory variables considered in the modeling process were fluorescence, transmissivity, oxygen concentration, salinity, region, year, season, diel period, depth bin, and volume of water filtered. The response variable was count, or number, of paralarvae.

Samples were assigned to one of six depth bins or removed from the dataset as described in the vertical distribution methods section above. The 1-meter binned CTD data were further binned to produce an average value for each variable over the actual open range of the net for each sample. Samples with missing CTD data were removed from the dataset. All variables were checked for outliers and samples containing an outlier in one or more explanatory variables were removed. Samples with outlying paralarvae count values were retained. Multicollinearity between explanatory variables was assessed using modified variance inflation factors (VIF) (DeRuiter 2019). Variables with a modified VIF above a threshold value of 4 were removed from the analysis (DeRuiter 2019).

The percentage of samples that had a value of zero for the response variable was calculated for all samples to evaluate whether the data were zero-inflated. Single variable and two variable tables were produced to explore sampling effort in time and space (Appendix 2). To

check for patterns in abundance due to multiple samples collected at same site, a boxplot was created of counts versus site ID.

Generalized additive models (GAMs) were created using three distributions appropriate for count data: Poisson, quasi-Poisson, and negative binomial. All variables selected by the VIF process were included in the model, with the continuous variables as modified thin plate regression spline smoothers (Miller et al. 2016). Volume filtered was tested to determine whether it was appropriate to include as an offset term or as a covariate (Ieno et al. 2014). Each model was refined by converting smooth terms to parametric terms if the smoother was linear. The dispersion statistic was calculated (Zuur and Camphuysen 2013). The model summary was produced and the deviance explained was recorded. For the negative binomial model, the theta value ($\theta$) was extracted from the model summary.

The model with the dispersion statistic closest to 1, indicating it was neither over- nor under-dispersed, and a positive adjusted $R^2$ value was selected for further model fitting. Backward stepwise selection was used to obtain the best-fitting model (Appendix 2). In this procedure, the variable with the highest p-value in the model was removed and the model was re-run. The two models (one with the variable and one without) were compared using a multi-model ANOVA and the Akaike Information Criterion (AIC). If either test indicated that the reduced model was better, the variable with the new highest p-value was removed from that model and it was re-run. The process was repeated until no variables had insignificant p-values ($\alpha = 0.05$). At this point, the working model was considered the best model. The estimates of the parametric coefficients were extracted from the model summary. The smooth functions and the partial effects of the parametric model components were plotted for interpretation.

Model validation was conducted to verify whether the final model met the assumptions of the technique and could be used for a robust interpretation. Homogeneity of variance was verified by plotting the Pearson residuals versus the fitted values. Model misfit (or independence) was verified by plotting the Pearson residuals versus each covariate in the

model and each covariate not included in the model. Normality was verified by assessing a histogram and Q-Q plot of the Pearson residuals. Influential observations were identified using Cook's distance with three thresholds (Appendix 2). To determine whether there were site-specific patterns that were not captured in the model, a boxplot was created of the Pearson residuals versus site. Finally, independence between samples collected at multiple spatial locations was verified by calculating Moran's I statistic using both the Pearson residuals and the original count data. Additionally, spatial correlation between samples collected at the same site versus at other sites was tested using Moran's I.

*2.4.4. Community Structure*

The beta-diversity (family composition and abundance) of the community was analyzed. The effects of the categorical variables year, season, diel period, region, and depth bin were examined. All samples containing at least one cephalopod of an identified taxon were used. To test whether beta-diversity differed by depth bin, the same subset of samples included in the analysis of vertical distribution was used. Counts were converted to abundance by dividing by the volume of water filtered by the sample and multiplying by 1000. Principal coordinates analysis (PCOA) was used to visualize similarities between samples based on their beta-diversity. It was also used to determine which taxa were more likely to be found together and which taxa drove the patterns in beta-diversity.

PCOA ordination was also used to visualize the dispersion and mean beta-diversity of samples for each of the variables listed above by plotting covariance ellipses and the centroid of each ellipse using the samples in each level of the grouping factor. The shape and size of these ellipses gives an idea of whether the dispersion of different groups varies while the location of the centroid gives an idea of whether the mean beta-diversity of the groups varies. All PCOA ordinations were created using non-transformed abundance data.

To determine if significant differences existed in the dispersion of samples among different levels of the grouping factors, which could also indicate changes in the variability of the beta-diversity, the PERMDISP algorithm was used, followed by a permutation-based multivariate analysis of variance (PERMANOVA) to determine if the dispersions calculated by PERMDISP were significantly different. This tests the null hypothesis of no significant difference in multivariate dispersion in the beta-diversity of paralarvae across observations grouped by the grouping factor under consideration ($\alpha = 0.05$). If the PERMANOVA was significant, a pairwise PERMANOVA was performed to determine which pairs of levels of the grouping factor were significantly different in dispersion.

Next, a PERMANOVA was run on the samples to determine if the grouping factor had a significant effect on the beta-diversity of the community. This tests the null hypothesis that there is no significant difference in the family composition and abundance (beta-diversity) of paralarvae among the levels of the grouping factor under consideration. A significant PERMANOVA following a significant PERMDISP means that the groups of samples differ in dispersion and potentially also in location. A significant PERMANOVA following a non-significant PERMDISP means that the groups of samples do not differ in dispersion, but do differ in location.

If the PERMANOVA on samples' taxa abundances was significant, a pairwise PERMANOVA was performed to identify which pairs of levels of the grouping factor were significantly different in beta-diversity. This tests the null hypothesis that there is no significant difference in the family composition and abundance of paralarvae among each pair of levels of the grouping factor. The PERMDISPs and PERMANOVAs were performed on square-root transformed abundance (# paralarvae per 1000 $m^3$) data, as this transformation down-weighs the importance of the more abundant taxa and gives more value to rare taxa. PERMANOVAs produce an F-statistic that describes how well the model captures the variability in the multivariate data compared to what remains in the error portion, where a larger F-statistic

means the model performs better. The PCOA ordinations can illustrate the differences in dispersion and location.

Finally, a canonical analysis of principal coordinates (CAP) was performed on the square-root transformed abundance data using each grouping factor to visualize the differences between groups of samples (based on grouping factor) and determine which levels of the grouping factor are the most similar. Axes on the CAP ordination maximize the separation between groups.

Indicator power values for each taxon were calculated based on each grouping factor using square-root transformed abundance data and taxa with significant indicator power values were plotted on the CAP ordinations to show which taxa characterize specific groups of the grouping factor, if any. Indicator power values greater than 0.90 are very strong, indicating a taxon has high group specificity (i.e., individuals of the taxon are mainly found in one group) and group fidelity (i.e., individuals of the taxon are found in a majority of the samples belonging to one group). Values less than 0.50 indicate weak group specificity, group fidelity, or both, and values less than 0.25 suggest that the taxon is not a good representative of a group, but is still representative, if the indicator power value was significant.

Additionally, CAP produces a classification model and tests the null hypothesis of no significant difference between the CAP-based model's classification success rate of paralarvae community samples based on grouping factor versus one produced by a classifier using a random allocation model. The classification success rate shows how well the CAP-based model identifies which group a sample came from. Identifying which other groups samples in a group were misidentified as belonging to can indicate similarity between the groups.

## 3. Results

Over the entire sampling period, 2240 paralarvae were examined from 637 samples. Sampling effort varied widely across time and space (Table 2, Table 3), based on opportunistic sampling on dedicated plankton surveys.

Table 2. Sampling effort by year and season.

| | *2009* | *2011* | *2012* | *Total* |
|---|---|---|---|---|
| *Winter* | 246 | 0 | 94 | **340** |
| *Spring* | 0 | 0 | 115 | **115** |
| *Early Fall* | 61 | 60 | 61 | **182** |
| *Total* | **307** | **60** | **270** | **637** |

Table 3. Sampling effort by season and region.

| | *NW GOM Shelf* | *NC GOM Shelf* | *WFS* | *GOM Slope* | *GOM Basin* | *Total* |
|---|---|---|---|---|---|---|
| *Winter* | 33 | 28 | 131 | 142 | 6 | **340** |
| *Spring* | 0 | 0 | 15 | 65 | 35 | **115** |
| *Early Fall* | 56 | 32 | 74 | 20 | 0 | **182** |
| *Total* | **89** | **60** | **220** | **227** | **41** | **637** |

Specimens were identified to the lowest possible taxonomic level, typically family, although genus and species level identifications were recorded where possible. Twenty-one families (1569 individuals) were identified. The families with the highest counts were Octopodidae (383 paralarvae), Enoploteuthidae (358), Ommastrephidae (335), Pyroteuthidae (192), and Loliginidae (140) (Figure 4). More than ten individuals were also collected and identified for the following taxa: Onychoteuthidae, Cranchiidae, Lycoteuthidae, Octopoteuthidae, and Argonautidae. The remaining identified taxa were Ancistrocheiridae, Chiroteuthidae, Sepiolidae, Cycloteuthidae, Joubiniteuthidae, Mastigoteuthidae, Thysanoteuthidae, and the singletons: Brachioteuthidae, Chtenopterygidae, Histioteuthidae, and Neoteuthidae. The remaining 671 paralarvae, or 30% of the total, could not be identified to family. These specimens were mostly very small (e.g., 1.2 mm mantle length) and/or damaged, so

distinguishing features had not yet been developed or were unable to be used for identification. They were identified as ‘unidentified squid’ for specimens that were clearly squid but could not be identified to a family or as ‘unidentified cephalopod’ for specimens that could not be identified to either squid or octopus.

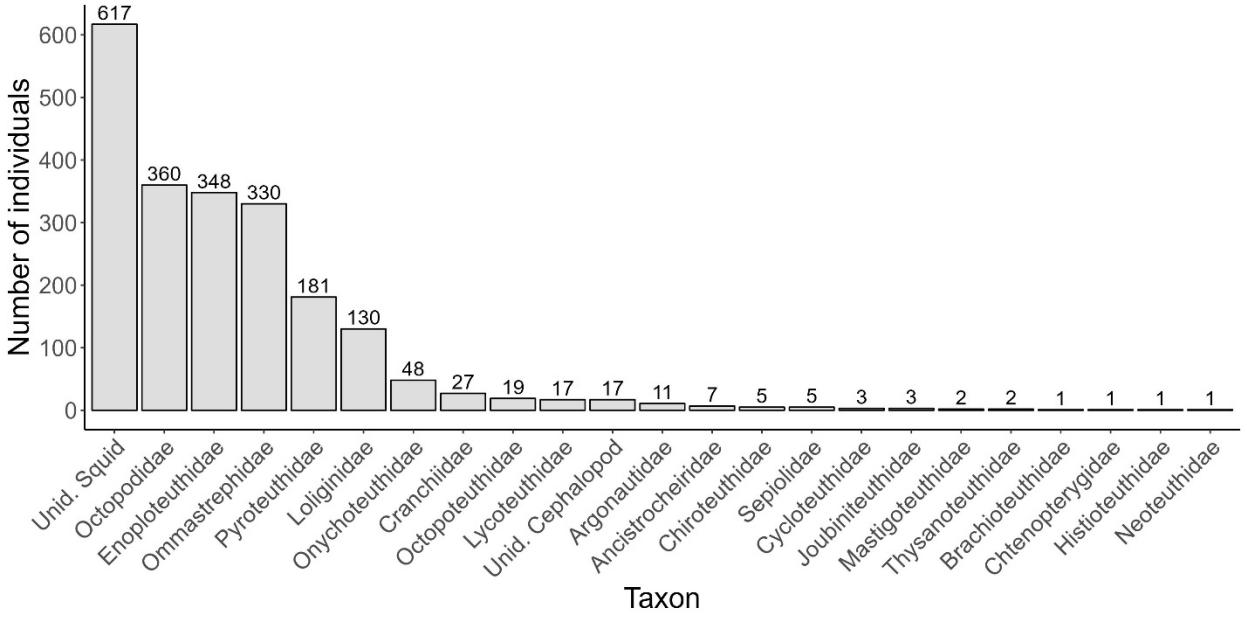


Figure 4. Number of individuals of each taxon collected from the six surveys, along with the individuals that could not be identified to family (Unid. Squid and Unid. Cephalopod).

3.1. Abundance and Distribution

Over the entire sampling period (six surveys), the mean abundance of all paralarvae was 21.9 paralarvae 1000 $m^{-3}$ (Table 4). The taxa with the highest mean abundance were Loliginidae (14.3 paralarvae 1000 $m^{-3}$), Octopodidae (10.5 paralarvae 1000 $m^{-3}$), Brachioteuthidae (9.80 paralarvae 1000 $m^{-3}$), Enoploteuthidae (9.35 paralarvae 1000 $m^{-3}$), and Sepiolidae (7.79 paralarvae 1000 $m^{-3}$). However, only one individual of the family Brachioteuthidae was collected over all cruises, a frequency of occurrence of 0.157% (Table 5). Similarly, only five Sepiolidae individuals were collected; the family occurred in 0.314% of samples collected over all cruises (Table 5). The frequency of occurrence of the other high abundance taxa ranged from 7% to

23% (Loliginidae: 7.06%, Octopodidae: 22.6%, Enoploteuthidae: 23.1%). Other taxa with high abundance and high frequency of occurrence were Ommastrephidae (7.68 paralarvae 1000 $m^{-3}$, 24.5% of samples) and Pyroteuthidae (6.25 paralarvae 1000 $m^{-3}$, 17.9% of samples).

Taxa with significant differences in mean abundance between years were Loliginidae, Lycoteuthidae, Octopodidae, Pyroteuthidae, and the group of all paralarvae. Pairs of years with significant differences in abundance are given in Table 6. Only Loliginidae and Octopodidae showed significant differences in mean abundance between seasons, along with the group of all paralarvae. Pairs of seasons with significant differences in abundance are given in Table 7. To omit any effect of year on abundance, abundance was also compared between seasons within the same year. In 2009, only winter and early fall were surveyed. Three taxa (Enoploteuthidae, Loliginidae, and Octopodidae), as well as the group of all paralarvae, had significantly different mean abundances between seasons (Table 8). In 2012, three seasons were surveyed. Only the group of all paralarvae showed significant differences in mean abundance between seasons. Pairs of seasons with significant differences in abundance are given in Table 9.

Only five taxa were identified in all five regions (northwestern GOM shelf, north-central GOM shelf, West Florida Shelf, GOM slope, and GOM basin): Cranchiidae, Enoploteuthidae, Octopodidae, Ommastrephidae, and Pyroteuthidae (Table 10). No taxa showed significant differences in mean abundance between regions. However, there were some qualitative differences in frequency of occurrence among regions for some taxa (Table 11). For example, Pyroteuthids were identified in 27.3% of slope samples, but fewer than 15% of samples in all other regions. Octopods were identified in 23.3-34.5% of shelf samples, but only 10% or less of slope and basin samples, respectively. Similarly, Loliginids were identified in 9.09-21.7% of shelf samples but fewer than 1% of slope and basin samples.

Table 4. Mean abundance and standard deviation of each cephalopod taxon in different sampling periods. The number of individuals of each taxon collected in each sampling period is listed in parentheses. These values were calculated from 637 samples.

| | *All Years* | *2009* | | | *2011* | *2012* | | | |
|---|---|---|---|---|---|---|---|---|---|
| *Taxon* | All Seasons | All | Winter | Early Fall | Early Fall | All | Winter | Spring | Early Fall |
| *All Paralarvae* | 21.89 ± 31.19 (2240) | 20.29 ± 20.93 (953) | 18.58 ± 15.3 (834) | 30.55 ± 39.86 (119) | 24.86 ± 32.08 (291) | 23.32 ± 41.76 (996) | 8.86 ± 6.06 (209) | 24.45 ± 22.6 (220) | 46.7 ± 67.58 (567) |
| *Ancistrocheiridae* | 3.24 ± 0.65 (7) | 4.39 ± 0 (1) | 4.39 ± 0 (1) | 0 ± 0 (0) | 2.61 ± 0 (1) | 3.14 ± 0.43 (5) | 0 ± 0 (0) | 2.85 ± 0.24 (3) | 3.57 ± 0.02 (2) |
| *Argonautidae* | 3.61 ± 0.58 (11) | 3.71 ± 0.48 (10) | 3.71 ± 0.48 (10) | 0 ± 0 (0) | 0 ± 0 (0) | 2.54 ± 0 (1) | 0 ± 0 (0) | 2.54 ± 0 (1) | 0 ± 0 (0) |
| *Brachioteuthidae* | 9.8 ± 0 (1) | 9.8 ± 0 (1) | 9.8 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) |
| *Chiroteuthidae* | 4.09 ± 1.64 (6) | 4.34 ± 1.78 (5) | 4.34 ± 1.78 (5) | 0 ± 0 (0) | 0 ± 0 (0) | 3.1 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) | 3.1 ± 0 (1) |
| *Chtenopterygidae* | 2.26 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 2.26 ± 0 (1) | 2.26 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) |
| *Cranchiidae* | 4.14 ± 1.6 (27) | 3.96 ± 1.25 (14) | 3.96 ± 1.3 (13) | 3.98 ± 0 (1) | 3.1 ± 0 (1) | 4.57 ± 2.14 (12) | 4.09 ± 1.83 (8) | 8.2 ± 0 (3) | 3.8 ± 0 (1) |
| *Cycloteuthidae* | 2.18 ± 0.25 (3) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 2.18 ± 0.25 (3) | 2.32 ± 0.09 (2) | 1.91 ± 0 (1) | 0 ± 0 (0) |
| *Enoploteuthidae* | 9.35 ± 20.24 (358) | 7.24 ± 5.37 (128) | 6.47 ± 4.21 (108) | 12.02 ± 8.78 (20) | 5.75 ± 3.67 (26) | 13.81 ± 33.39 (204) | 4.88 ± 2.56 (30) | 6.16 ± 3.94 (32) | 32.05 ± 56.35 (142) |
| *Histioteuthidae* | 2.94 ± 0 (1) | 2.94 ± 0 (1) | 2.94 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) |
| *Joubiniteuthidae* | 2.87 ± 1.2 (3) | 3.18 ± 1.53 (2) | 3.18 ± 1.53 (2) | 0 ± 0 (0) | 0 ± 0 (0) | 2.26 ± 0 (1) | 2.26 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) |
| *Loliginidae* | 14.35 ± 29.1 (140) | 13.71 ± 12.67 (34) | 4.71 ± 1.77 (10) | 20.26 ± 13.22 (24) | 19.28 ± 44.08 (93) | 4.75 ± 4.74 (13) | 4.75 ± 4.74 (13) | 0 ± 0 (0) | 0 ± 0 (0) |
| *Lycoteuthidae* | 4.76 ± 2.64 (21) | 5.76 ± 2.14 (12) | 5.87 ± 2.26 (11) | 4.9 ± 0 (1) | 6 ± 6.33 (3) | 2.84 ± 0.46 (6) | 2.25 ± 0 (1) | 2.77 ± 0.23 (2) | 3.09 ± 0.47 (3) |
| *Mastigoteuthidae* | 3.56 ± 1.11 (2) | 4.35 ± 0 (1) | 4.35 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) | 2.78 ± 0 (1) | 2.78 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) |
| *Neoteuthidae* | 3.46 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 3.46 ± 0 (1) | 0 ± 0 (0) | 3.46 ± 0 (1) | 0 ± 0 (0) |
| *Octopodidae* | 10.51 ± 17.3 (383) | 11.01 ± 23.72 (126) | 7.13 ± 4.37 (95) | 36.22 ± 61.38 (31) | 13.3 ± 12.14 (124) | 8.23 ± 9.4 (133) | 5.6 ± 4.04 (36) | 6.11 ± 2.54 (18) | 10.9 ± 12.62 (79) |
| *Octopoteuthidae* | 4.47 ± 4.19 (20) | 3.15 ± 0.87 (4) | 3.15 ± 0.87 (4) | 0 ± 0 (0) | 0 ± 0 (0) | 4.95 ± 4.83 (16) | 3.94 ± 0 (1) | 4.88 ± 2.75 (3) | 5.09 ± 5.67 (12) |
| *Ommastrephidae* | 7.68 ± 7.67 (335) | 7.86 ± 7.87 (170) | 7.81 ± 8.19 (157) | 8.35 ± 4.46 (13) | 7.24 ± 7.7 (23) | 7.48 ± 7.49 (142) | 4.61 ± 3.01 (33) | 10.91 ± 11.29 (54) | 8.03 ± 6.56 (55) |
| *Onychoteuthidae* | 4.56 ± 3.46 (50) | 4.4 ± 2.37 (20) | 4.4 ± 2.37 (20) | 0 ± 0 (0) | 0 ± 0 (0) | 4.73 ± 4.36 (30) | 3.69 ± 1.99 (11) | 3.21 ± 0.39 (2) | 6.79 ± 6.83 (17) |
| *Pyroteuthidae* | 6.25 ± 4.86 (192) | 6.86 ± 5.12 (169) | 6.85 ± 5.14 (167) | 7.91 ± 0 (2) | 0 ± 0 (0) | 3.38 ± 1.28 (23) | 3.01 ± 0.79 (11) | 2.61 ± 0 (1) | 3.86 ± 1.64 (11) |
| *Sepiolidae* | 7.79 ± 5.62 (5) | 11.76 ± 0 (4) | 11.76 ± 0 (4) | 0 ± 0 (0) | 0 ± 0 (0) | 3.82 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) | 3.82 ± 0 (1) |
| *Thysanoteuthidae* | 3.65 ± 0.04 (2) | 3.65 ± 0.04 (2) | 3.65 ± 0.04 (2) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) |

Table 5. Percent of all paralarvae collected in each sampling period and frequency of occurrence in each sampling period. (% of All – percent of all paralarvae; % FO – percent frequency of occurrence)

| | *All Years* | | *2009* | | | | | | *2011* | | *2012* | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Taxon* | All Seasons | | All | | Winter | | Early Fall | | Early Fall | | All | | Winter | | Spring | | Early Fall | |
| | % of All | % FO | % of All | % FO | % of All | % FO | % of All | % FO | % of All | % FO | % of All | % FO | % of All | % FO | % of All | % FO | % of All | % FO |
| *All Paralarvae* | 100 | 60.28 | 100 | 66.12 | 100 | 70.73 | 100 | 47.54 | 100 | 70 | 100 | 51.48 | 100 | 74.47 | 100 | 23.48 | 100 | 68.85 |
| *Ancistrocheiridae* | 0.31 | 1.1 | 0.1 | 0.33 | 0.12 | 0.41 | 0 | 0 | 0.34 | 1.67 | 0.5 | 1.85 | 0 | 0 | 1.36 | 2.61 | 0.35 | 3.28 |
| *Argonautidae* | 0.49 | 1.73 | 1.05 | 3.26 | 1.2 | 4.07 | 0 | 0 | 0 | 0 | 0.1 | 0.37 | 0 | 0 | 0.45 | 0.87 | 0 | 0 |
| *Brachioteuthidae* | 0.04 | 0.16 | 0.1 | 0.33 | 0.12 | 0.41 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Chiroteuthidae* | 0.27 | 0.78 | 0.52 | 1.3 | 0.6 | 1.63 | 0 | 0 | 0 | 0 | 0.1 | 0.37 | 0 | 0 | 0 | 0 | 0.18 | 1.64 |
| *Chtenopterygidae* | 0.04 | 0.16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.1 | 0.37 | 0.48 | 1.06 | 0 | 0 | 0 | 0 |
| *Cranchiidae* | 1.21 | 3.45 | 1.47 | 4.23 | 1.56 | 4.88 | 0.84 | 1.64 | 0.34 | 1.67 | 1.2 | 2.96 | 3.83 | 6.38 | 1.36 | 0.87 | 0.18 | 1.64 |
| *Cycloteuthidae* | 0.13 | 0.47 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.3 | 1.11 | 0.96 | 2.13 | 0.45 | 0.87 | 0 | 0 |
| *Enoploteuthidae* | 15.98 | 23.08 | 13.43 | 25.73 | 12.95 | 27.64 | 16.81 | 18.03 | 8.93 | 28.33 | 20.48 | 18.89 | 14.35 | 20.21 | 14.55 | 13.91 | 25.04 | 26.23 |
| *Histioteuthidae* | 0.04 | 0.16 | 0.1 | 0.33 | 0.12 | 0.41 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Joubiniteuthidae* | 0.13 | 0.47 | 0.21 | 0.65 | 0.24 | 0.81 | 0 | 0 | 0 | 0 | 0.1 | 0.37 | 0.48 | 1.06 | 0 | 0 | 0 | 0 |
| *Loliginidae* | 6.25 | 7.06 | 3.57 | 6.19 | 1.2 | 3.25 | 20.17 | 18.03 | 31.96 | 30 | 1.31 | 2.96 | 6.22 | 8.51 | 0 | 0 | 0 | 0 |
| *Lycoteuthidae* | 0.94 | 2.67 | 1.26 | 2.93 | 1.32 | 3.25 | 0.84 | 1.64 | 1.03 | 3.33 | 0.6 | 2.22 | 0.48 | 1.06 | 0.91 | 1.74 | 0.53 | 4.92 |
| *Mastigoteuthidae* | 0.09 | 0.31 | 0.1 | 0.33 | 0.12 | 0.41 | 0 | 0 | 0 | 0 | 0.1 | 0.37 | 0.48 | 1.06 | 0 | 0 | 0 | 0 |
| *Neoteuthidae* | 0.04 | 0.16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.1 | 0.37 | 0 | 0 | 0.45 | 0.87 | 0 | 0 |
| *Octopodidae* | 17.1 | 22.61 | 13.22 | 19.54 | 11.39 | 21.14 | 26.05 | 13.11 | 42.61 | 53.33 | 13.35 | 19.26 | 17.22 | 20.21 | 8.18 | 6.96 | 13.93 | 40.98 |
| *Octopoteuthidae* | 0.89 | 2.35 | 0.42 | 1.3 | 0.48 | 1.63 | 0 | 0 | 0 | 0 | 1.61 | 4.07 | 0.48 | 1.06 | 1.36 | 1.74 | 2.12 | 13.11 |
| *Ommastrephidae* | 14.96 | 24.49 | 17.84 | 28.34 | 18.82 | 31.71 | 10.92 | 14.75 | 7.9 | 18.33 | 14.26 | 21.48 | 15.79 | 23.4 | 24.55 | 13.04 | 9.7 | 34.43 |
| *Onychoteuthidae* | 2.23 | 5.34 | 2.1 | 5.54 | 2.4 | 6.91 | 0 | 0 | 0 | 0 | 3.01 | 6.3 | 5.26 | 9.57 | 0.91 | 1.74 | 3 | 9.84 |
| *Pyroteuthidae* | 8.57 | 17.9 | 17.73 | 30.62 | 20.02 | 37.8 | 1.68 | 1.64 | 0 | 0 | 2.31 | 7.41 | 5.26 | 10.64 | 0.45 | 0.87 | 1.94 | 14.75 |
| *Sepiolidae* | 0.22 | 0.31 | 0.42 | 0.33 | 0.48 | 0.41 | 0 | 0 | 0 | 0 | 0.1 | 0.37 | 0 | 0 | 0 | 0 | 0.18 | 1.64 |
| *Thysanoteuthidae* | 0.09 | 0.31 | 0.21 | 0.65 | 0.24 | 0.81 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Table 6. Taxa with significant differences in mean abundance between years. Bold years had significantly higher abundance than the year they are being compared to. Significant p-values are in bold. Pairwise p-values are adjusted using Holm's method.

| *Taxon* | *p-value* | | | *Adj. p-value* | | | *Adj. p-value* | | | *Adj. p-value* |
|---|---|---|---|---|---|---|---|---|---|---|
| *All Paralarvae* | **0.006** | 2009 | 2011 | 1.000 | 2009 | **2012** | **0.006** | 2011 | 2012 | 0.105 |
| *Loliginidae* | **0.008** | 2009 | 2011 | 0.450 | **2009** | 2012 | **0.007** | **2011** | 2012 | **0.028** |
| *Lycoteuthidae* | **0.033** | 2009 | 2011 | 0.624 | **2009** | 2012 | **0.027** | 2011 | 2012 | 0.624 |
| *Octopodidae* | **0.032** | 2009 | 2011 | 0.229 | 2009 | 2012 | 0.199 | **2011** | 2012 | **0.031** |
| *Pyroteuthidae* | **4.289e-06** | 2009 | NA | NA | **2009** | 2012 | **4.213e-06** | NA | 2012 | NA |

Table 7. Taxa with significant differences in mean abundance between seasons. Bold seasons had significantly higher abundance than the season they are being compared to. Significant p-values are in bold. Pairwise p-values are adjusted using Holm's method.

| *Taxon* | *p-value* | | | *Adj. p-value* | | | *Adj. p-value* | | | *Adj. p-value* |
|---|---|---|---|---|---|---|---|---|---|---|
| *All Paralarvae* | **8.601e-05** | Winter | Spring | 0.128 | Winter | **Early Fall** | **8.854e-05** | Spring | Early Fall | 0.642 |
| *Loliginidae* | **2.134e-04** | Winter | Spring | NA | Winter | **Early Fall** | **3.756e-04** | Spring | Early Fall | NA |
| *Octopodidae* | **0.018** | Winter | Spring | 0.934 | Winter | **Early Fall** | **0.018** | Spring | Early Fall | 0.357 |

Table 8. Taxa with significant differences in mean abundance between seasons in 2009. Bold seasons had significantly higher abundance than the season they are being compared to. Significant p-values are in bold.

| *Taxon* | | | *p-value* |
|---|---|---|---|
| *All Paralarvae* | Winter | **Early Fall** | **0.017** |
| *Enoploteuthidae* | Winter | **Early Fall** | **0.002** |
| *Loliginidae* | Winter | **Early Fall** | **1.852e-04** |
| *Octopodidae* | Winter | **Early Fall** | **0.002** |

Table 9. Taxa with significant differences in mean abundance between seasons in 2012. Bold seasons had significantly higher abundance than the season they are being compared to. Significant p-values are in bold. Pairwise p-values are adjusted using Holm's method.

| *Taxon* | *p-value* | | | *Adj. p-value* | | | *Adj. p-value* | | | *Adj. p-value* |
|---|---|---|---|---|---|---|---|---|---|---|
| *All Paralarvae* | **1.004e-07** | Winter | **Spring** | **2.289e-04** | Winter | **Early Fall** | **6.380e-07** | Spring | Early Fall | 0.574 |

Table 10. Mean abundance and standard deviation of each cephalopod taxon in each sampling region. The number of individuals of each taxon collected in each sampling region is listed in parentheses. These values were calculated from 637 samples.

| *Taxon* | *NW GOM Shelf* | *NC GOM Shelf* | *West Florida Shelf* | *GOM Slope* | *GOM Basin* |
|---|---|---|---|---|---|
| *All Paralarvae* | 24.14 ± 31.18 (281) | 19.72 ± 15.46 (152) | 23.99 ± 42.64 (990) | 18.47 ± 16.9 (696) | 29.39 ± 24.01 (121) |
| *Ancistrocheiridae* | 2.61 ± 0 (1) | 0 ± 0 (0) | 3.57 ± 0.02 (2) | 3.62 ± 1.09 (2) | 2.85 ± 0.34 (2) |
| *Argonautidae* | 4.03 ± 0.44 (4) | 0 ± 0 (0) | 3.77 ± 0 (1) | 3.3 ± 0.54 (6) | 0 ± 0 (0) |
| *Brachioteuthidae* | 0 ± 0 (0) | 9.8 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) |
| *Chiroteuthidae* | 0 ± 0 (0) | 5.65 ± 0 (1) | 3.1 ± 0 (1) | 3.9 ± 1.9 (4) | 0 ± 0 (0) |
| *Chtenopterygidae* | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 2.26 ± 0 (1) | 0 ± 0 (0) |
| *Cranchiidae* | 3.93 ± 0.82 (3) | 3.1 ± 0 (1) | 3.72 ± 0.68 (5) | 4.44 ± 1.99 (17) | 3.61 ± 0 (1) |
| *Cycloteuthidae* | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 2.18 ± 0.25 (3) | 0 ± 0 (0) |
| *Enoploteuthidae* | 8.63 ± 6.52 (53) | 6.16 ± 3.76 (26) | 14.63 ± 34.69 (188) | 5.86 ± 4.13 (77) | 8.26 ± 5.32 (14) |
| *Histioteuthidae* | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 2.94 ± 0 (1) | 0 ± 0 (0) |
| *Joubiniteuthidae* | 0 ± 0 (0) | 0 ± 0 (0) | 2.26 ± 0 (1) | 3.18 ± 1.53 (2) | 0 ± 0 (0) |
| *Loliginidae* | 9.06 ± 8.63 (15) | 15.56 ± 13.63 (30) | 17.02 ± 42.12 (94) | 3.38 ± 0 (1) | 0 ± 0 (0) |
| *Lycoteuthidae* | 4.9 ± 0 (1) | 0 ± 0 (0) | 4.97 ± 2.92 (14) | 5.16 ± 2.97 (4) | 2.83 ± 0.32 (2) |
| *Mastigoteuthidae* | 0 ± 0 (0) | 2.78 ± 0 (1) | 0 ± 0 (0) | 4.35 ± 0 (1) | 0 ± 0 (0) |
| *Neoteuthidae* | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 3.46 ± 0 (1) | 0 ± 0 (0) |
| *Octopodidae* | 15.72 ± 34.02 (74) | 10.96 ± 10.59 (40) | 9.79 ± 9.82 (222) | 6.52 ± 4.62 (44) | 5.13 ± 2.89 (3) |
| *Octopoteuthidae* | 3.94 ± 0 (1) | 0 ± 0 (0) | 5.19 ± 5.63 (12) | 3.73 ± 1.91 (6) | 2.94 ± 0 (1) |
| *Ommastrephidae* | 8.43 ± 6.42 (49) | 9.16 ± 7.89 (23) | 7.03 ± 6.54 (115) | 6.59 ± 7.71 (101) | 14.4 ± 12.46 (47) |
| *Onychoteuthidae* | 4.64 ± 2.27 (5) | 3.2 ± 0.19 (2) | 5.66 ± 5.32 (21) | 4.09 ± 2.5 (22) | 0 ± 0 (0) |
| *Pyroteuthidae* | 8.04 ± 8.65 (25) | 7.82 ± 1.74 (10) | 4.61 ± 2.06 (38) | 6.66 ± 4.88 (116) | 3.21 ± 0.68 (3) |
| *Sepiolidae* | 0 ± 0 (0) | 0 ± 0 (0) | 3.82 ± 0 (1) | 11.76 ± 0 (4) | 0 ± 0 (0) |
| *Thysanoteuthidae* | 0 ± 0 (0) | 0 ± 0 (0) | 3.68 ± 0 (1) | 3.62 ± 0 (1) | 0 ± 0 (0) |

Table 11. Percent of all paralarvae collected in each sampling region and frequency of occurrence in each sampling region. (% FO – percent frequency of occurrence)

| | *NW GOM Shelf* | | *NC GOM Shelf* | | *West Florida Shelf* | | *GOM Slope* | | *GOM Basin* | |
|---|---|---|---|---|---|---|---|---|---|---|
| *Taxon* | % of All Paralarvae | % FO | % of All Paralarvae | % FO | % of All Paralarvae | % FO | % of All Paralarvae | % FO | % of All Paralarvae | % FO |
| *All Paralarvae* | 100 | 64.04 | 100 | 61.67 | 100 | 65.91 | 100 | 58.15 | 100 | 31.71 |
| *Ancistrocheiridae* | 0.36 | 1.12 | 0 | 0 | 0.2 | 0.91 | 0.29 | 0.88 | 1.65 | 4.88 |
| *Argonautidae* | 1.42 | 4.49 | 0 | 0 | 0.1 | 0.45 | 0.86 | 2.64 | 0 | 0 |
| *Brachioteuthidae* | 0 | 0 | 0.66 | 1.67 | 0 | 0 | 0 | 0 | 0 | 0 |
| *Chiroteuthidae* | 0 | 0 | 0.66 | 1.67 | 0.1 | 0.45 | 0.57 | 1.32 | 0 | 0 |
| *Chtenopterygidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0.14 | 0.44 | 0 | 0 |
| *Cranchiidae* | 1.07 | 3.37 | 0.66 | 1.67 | 0.51 | 1.82 | 2.44 | 5.73 | 0.83 | 2.44 |
| *Cycloteuthidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0.43 | 1.32 | 0 | 0 |
| *Enoploteuthidae* | 18.86 | 32.58 | 17.11 | 31.67 | 18.99 | 21.36 | 11.06 | 20.26 | 11.57 | 14.63 |
| *Histioteuthidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0.14 | 0.44 | 0 | 0 |
| *Joubiniteuthidae* | 0 | 0 | 0 | 0 | 0.1 | 0.45 | 0.29 | 0.88 | 0 | 0 |
| *Loliginidae* | 5.34 | 12.36 | 19.74 | 21.67 | 9.49 | 9.09 | 0.14 | 0.44 | 0 | 0 |
| *Lycoteuthidae* | 0.36 | 1.12 | 0 | 0 | 1.41 | 4.55 | 0.57 | 1.76 | 1.65 | 4.88 |
| *Mastigoteuthidae* | 0 | 0 | 0.66 | 1.67 | 0 | 0 | 0.14 | 0.44 | 0 | 0 |
| *Neoteuthidae* | 0 | 0 | 0 | 0 | 0 | 0 | 0.14 | 0.44 | 0 | 0 |
| *Octopodidae* | 26.33 | 32.58 | 26.32 | 23.33 | 22.42 | 34.55 | 6.32 | 10.13 | 2.48 | 4.88 |
| *Octopoteuthidae* | 0.36 | 1.12 | 0 | 0 | 1.21 | 3.64 | 0.86 | 2.2 | 0.83 | 2.44 |
| *Ommastrephidae* | 17.44 | 24.72 | 15.13 | 16.67 | 11.62 | 25.91 | 14.51 | 25.11 | 38.84 | 24.39 |
| *Onychoteuthidae* | 1.78 | 4.49 | 1.32 | 3.33 | 2.12 | 4.55 | 3.16 | 7.93 | 0 | 0 |
| *Pyroteuthidae* | 8.9 | 14.61 | 6.58 | 10 | 3.84 | 13.64 | 16.67 | 27.31 | 2.48 | 7.32 |
| *Sepiolidae* | 0 | 0 | 0 | 0 | 0.1 | 0.45 | 0.57 | 0.44 | 0 | 0 |
| *Thysanoteuthidae* | 0 | 0 | 0 | 0 | 0.1 | 0.45 | 0.14 | 0.44 | 0 | 0 |

3.2. Vertical Distribution

The most frequently sampled depth bins in order of frequency were 0-20 m, 20-40 m, 40-60 m, 60-80 m, 80-100 m, 0-10 m, 20-30 m, and 100-130 m. The following depth bins were selected for analysis: 0-20 m, 20-40 m, 40-60 m, 60-80 m, 80-100 m, and 100-130 m. Of 637 original samples, 607 fit within one of the six depth bins and were retained for subsequent analysis. In this subset of samples, 2136 paralarvae belonging to 21 taxa and two unidentified groups were collected. The most abundant taxa (Octopodidae, Enoploteuthidae, and Ommastrephidae) had over 300 individuals collected while the least abundant taxa (Brachioteuthidae, Chtenopterygidae, Histioteuthidae, and Neoteuthidae) were only collected once.

Considering diel period, five taxa were only collected at one time of day. Nine families had a higher mean abundance in daytime samples while seven taxa and the group of all paralarvae had a higher mean abundance in nighttime samples (Table 12). Only two taxa had significantly difference abundances between day and night, Loliginidae and Pyroteuthidae. Both were significantly more abundant in daytime samples. Standard deviations were large for four of the five highest total count taxa, Enoploteuthidae, Loliginidae, Octopodidae, Ommastrephidae, as well as the group of all paralarvae, indicating high variability in abundance.

Table 12. Mean depth-aggregated abundance and standard deviation of each cephalopod taxon in daytime and nighttime samples. The number of individuals of each taxon collected in each time of day is listed in parentheses. Significant p-values are in bold. These values were calculated from 607 samples.

| *Taxon* | *Day* | *Night* | *p-value* |
|---|---|---|---|
| *All Paralarvae* | 21.26 ± 26.36 (996) | 22.88 ± 36.38 (1140) | 0.533 |
| *Ancistrocheiridae* | 2.85 ± 0.34 (2) | 3.4 ± 0.7 (5) | 0.434 |
| *Argonautidae* | 3.55 ± 0.48 (4) | 3.64 ± 0.66 (7) | 0.927 |
| *Brachioteuthidae* | 0 ± 0 (0) | 9.8 ± 0 (1) | only collected at one time of day |
| *Chiroteuthidae* | 2.41 ± 0 (1) | 4.13 ± 1.66 (4) | 0.5 |
| *Chtenopterygidae* | 0 ± 0 (0) | 2.26 ± 0 (1) | only collected at one time of day |
| *Cranchiidae* | 4.33 ± 1.62 (10) | 4.03 ± 1.64 (17) | 0.413 |
| *Cycloteuthidae* | 2.15 ± 0.34 (2) | 2.26 ± 0 (1) | 1 |
| *Enoploteuthidae* | 6.83 ± 5.45 (111) | 11.87 ± 27.72 (237) | 0.799 |
| *Histioteuthidae* | 0 ± 0 (0) | 2.94 ± 0 (1) | only collected at one time of day |
| *Joubiniteuthidae* | 3.26 ± 1.41 (2) | 2.1 ± 0 (1) | 0.667 |
| *Loliginidae* | 26.34 ± 44.93 (94) | 7.06 ± 6.61 (36) | **0.015** |
| *Lycoteuthidae* | 2.63 ± 1.12 (3) | 4.58 ± 1.88 (14) | 0.126 |
| *Mastigoteuthidae* | 4.35 ± 0 (1) | 2.78 ± 0 (1) | 1 |
| *Neoteuthidae* | 3.46 ± 0 (1) | 0 ± 0 (0) | only collected at one time of day |
| *Octopodidae* | 11.24 ± 21.92 (202) | 9.91 ± 10.57 (158) | 0.94 |
| *Octopoteuthidae* | 3.87 ± 1.64 (7) | 5.15 ± 5.64 (12) | 0.95 |
| *Ommastrephidae* | 9.17 ± 10.05 (175) | 6.65 ± 4.98 (155) | 0.468 |
| *Onychoteuthidae* | 5.68 ± 4.66 (30) | 3.65 ± 1.42 (18) | 0.061 |
| *Pyroteuthidae* | 7.46 ± 4.73 (93) | 5.39 ± 5.04 (88) | **0.001** |
| *Sepiolidae* | 0 ± 0 (0) | 7.79 ± 5.62 (5) | only collected at one time of day |
| *Thysanoteuthidae* | 3.68 ± 0 (1) | 3.62 ± 0 (1) | 1 |

Considering depth aggregated over time of day, four taxa were only collected in one depth bin (Table 13). Seven of the remaining taxa and the group of all paralarvae were collected in all six depth bins. Only two taxa, Octopodidae and Ommastrephidae, as well as the group of all paralarvae, had significantly different mean abundances between depth bins (Table 13). However, none of the pairs of depth bins were significantly different for Octopodidae considering the Holm's adjusted p-values. For Ommastrephidae, the mean abundance in the 0-20 m depth bin was significantly higher than that in the 60-80 m depth bin and the mean abundance in the 20-40 m depth bin was also significantly higher than the 60-80 m depth bin. For the group all paralarvae, mean abundance was significantly higher in the 0-20 m depth bin compared to the 80-100 m depth bin. The mean abundance of the 20-40 m depth bin was significantly higher than the 60-80 m, 80-100 m, and 100-130 m depth bins. The mean abundance of the 40-60 m depth bin was also significantly higher than the three deeper bins.

Considering depth during the day, four taxa were not collected in daytime samples and another four were collected in only one depth bin during the day (Table 14). Only four taxa and the group of all paralarvae were collected in every depth bin during the day. Comparing mean abundances between depth bins for daytime samples, only the group of all paralarvae had a significant difference in mean abundance. However, none of the pairs of bins had significant Holm's adjusted p-values.

Table 13. Mean time-of-day-aggregated abundance and standard deviation of each cephalopod taxon in each depth bin. The number of individuals of each taxon collected in each depth bin is listed in parentheses. Significant p-values are in bold. These values were calculated from 607 samples.

| *Taxon* | *0-20 m* | *20-40 m* | *40-60 m* | *60-80 m* | *80-100 m* | *100-130* | *p-value* |
|---|---|---|---|---|---|---|---|
| *All Paralarvae* | 24.09 ± 25.99 (448) | 29.29 ± 4.23 (770) | 26.19 ± 38.87 (511) | 12.29 ± 7.81 (177) | 11.53 ± 8.9 (126) | 11.51 ± 8.83 (104) | **6.19E-06** |
| *Ancistrocheiridae* | 2.85 ± 0 (1) | 3.36 ± 0.92 (3) | 3.08 ± 0.67 (2) | 3.58 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) | 0.716 |
| *Argonautidae* | 4.13 ± 0.36 (4) | 3.41 ± 0.59 (4) | 0 ± 0 (0) | 3.18 ± 0.28 (3) | 0 ± 0 (0) | 0 ± 0 (0) | 0.052 |
| *Brachioteuthidae* | 9.8 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | only collected in one depth bin |
| *Chiroteuthidae* | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 3.25 ± 0 (1) | 3.1 ± 0 (1) | 4.23 ± 2.57 (3) | 0.861 |
| *Chtenopterygidae* | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 2.26 ± 0 (1) | only collected in one depth bin |
| *Cranchiidae* | 4.02 ± 0 (1) | 4.29 ± 1.95 (8) | 6.35 ± 2.24 (5) | 3.34 ± 0.63 (2) | 3.6 ± 0.55 (8) | 3.4 ± 1.62 (3) | 0.361 |
| *Cycloteuthidae* | 1.91 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) | 2.39 ± 0 (1) | 0 ± 0 (0) | 2.26 ± 0 (1) | 0.368 |
| *Enoploteuthidae* | 9.72 ± 14.4 (87) | 13.79 ± 33.54 (163) | 6.94 ± 4.6 (59) | 5.29 ± 3.69 (26) | 6.34 ± 4.92 (9) | 4.92 ± 1.95 (4) | 0.729 |
| *Histioteuthidae* | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 2.94 ± 0 (1) | only collected in one depth bin |
| *Joubiniteuthidae* | 0 ± 0 (0) | 4.26 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 2.18 ± 0.11 (2) | 0.667 |
| *Loliginidae* | 5.91 ± 2.72 (13) | 14.2 ± 14.09 (41) | 28.39 ± 58.54 (68) | 12.06 ± 6.05 (7) | 2.84 ± 0 (1) | 0 ± 0 (0) | 0.544 |
| *Lycoteuthidae* | 3.68 ± 1.06 (2) | 6.8 ± 0 (2) | 4.54 ± 1.94 (5) | 4.03 ± 2.67 (5) | 3.79 ± 0.98 (2) | 2.25 ± 0 (1) | 0.652 |
| *Mastigoteuthidae* | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 4.35 ± 0 (1) | 2.78 ± 0 (1) | 0 ± 0 (0) | 1 |
| *Neoteuthidae* | 3.46 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | only collected in one depth bin |
| *Octopodidae* | 10.4 ± 9.98 (71) | 11.3 ± 10.94 (133) | 16.81 ± 35.91 (94) | 5.27 ± 3.23 (30) | 6.13 ± 4.23 (22) | 5.45 ± 3.65 (10) | **0.049** |
| *Octopoteuthidae* | 3.06 ± 0.17 (2) | 6.37 ± 6.42 (11) | 3.91 ± 0.33 (3) | 0 ± 0 (0) | 3.26 ± 0 (1) | 2.56 ± 0.65 (2) | 0.43 |
| *Ommastrephidae* | 9.21 ± 8.66 (83) | 8.94 ± 9.25 (117) | 6.76 ± 4.75 (75) | 3.67 ± 2.06 (16) | 5.01 ± 2.98 (20) | 15.01 ± 15.61 (19) | **0.013** |
| *Onychoteuthidae* | 4.22 ± 2.17 (5) | 6.79 ± 6.34 (19) | 4.46 ± 1.76 (10) | 3.24 ± 0.89 (7) | 4.58 ± 0 (2) | 3.79 ± 2.1 (5) | 0.524 |
| *Pyroteuthidae* | 10.12 ± 9.46 (24) | 5.48 ± 3.06 (29) | 6.16 ± 4.42 (42) | 6.27 ± 4.41 (34) | 6.15 ± 4.54 (28) | 4.95 ± 3.75 (24) | 0.104 |
| *Sepiolidae* | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 3.82 ± 0 (1) | 0 ± 0 (0) | 11.76 ± 0 (4) | 1 |
| *Thysanoteuthidae* | 3.62 ± 0 (1) | 3.68 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 1 |

Table 14. Mean abundance and standard deviation of each cephalopod taxon in each depth bin in daytime samples. The number of individuals of each taxon collected in each depth bin in daytime samples is listed in parentheses. Significant p-values are in bold. These values were calculated from 607 samples.

| *Taxon* | *0-20 m* | *20-40 m* | *40-60 m* | *60-80 m* | *80-100 m* | *100-130 m* | *p-value* |
|---|---|---|---|---|---|---|---|
| *All Paralarvae* | 18.45 ± 18.55 (155) | 25.1 ± 21.84 (334) | 30.94 ± 44.8 (293) | 12.74 ± 7.42 (97) | 13.01 ± 9.55 (66) | 12.3 ± 10.61 (51) | **0.006** |
| *Ancistrocheiridae* | 0 ± 0 (0) | 3.09 ± 0 (1) | 2.61 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 1 |
| *Argonautidae* | 4.03 ± 0 (1) | 3.77 ± 0 (1) | 0 ± 0 (0) | 3.19 ± 0.4 (2) | 0 ± 0 (0) | 0 ± 0 (0) | 0.259 |
| *Brachioteuthidae* | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | not collected at this time of day |
| *Chiroteuthidae* | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 2.41 ± 0 (1) | only collected in one depth bin |
| *Chtenopterygidae* | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | not collected at this time of day |
| *Cranchiidae* | 0 ± 0 (0) | 6.08 ± 2.99 (4) | 3.8 ± 0 (1) | 0 ± 0 (0) | 3.74 ± 0.56 (5) | 0 ± 0 (0) | 0.209 |
| *Cycloteuthidae* | 1.91 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) | 2.39 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) | 1 |
| *Enoploteuthidae* | 5.14 ± 3.54 (15) | 8.27 ± 7.44 (40) | 6.94 ± 4.68 (32) | 4.78 ± 1.98 (15) | 8.46 ± 5.83 (6) | 5.52 ± 2.33 (3) | 0.313 |
| *Histioteuthidae* | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | not collected at this time of day |
| *Joubiniteuthidae* | 0 ± 0 (0) | 4.26 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 2.26 ± 0 (1) | 1 |
| *Loliginidae* | 5.19 ± 3.13 (3) | 17.76 ± 15.32 (31) | 63.65 ± 86.66 (60) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0.145 |
| *Lycoteuthidae* | 0 ± 0 (0) | 0 ± 0 (0) | 2.61 ± 0 (1) | 2.64 ± 1.58 (2) | 0 ± 0 (0) | 0 ± 0 (0) | 1 |
| *Mastigoteuthidae* | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 4.35 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) | only collected in one depth bin |
| *Neoteuthidae* | 3.46 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | only collected in one depth bin |
| *Octopodidae* | 10.81 ± 9.81 (15) | 9.38 ± 7.93 (60) | 22.48 ± 43.72 (78) | 5.35 ± 2.6 (20) | 6.88 ± 4.56 (19) | 5.45 ± 3.65 (10) | 0.208 |
| *Octopoteuthidae* | 2.94 ± 0 (1) | 4.43 ± 3.39 (3) | 4.08 ± 0.2 (2) | 0 ± 0 (0) | 3.26 ± 0 (1) | 0 ± 0 (0) | 0.733 |
| *Ommastrephidae* | 11.96 ± 12.12 (46) | 10.81 ± 11.76 (68) | 6.49 ± 4.18 (33) | 4.02 ± 2.93 (9) | 6.66 ± 4.58 (8) | 20.6 ± 25.68 (11) | 0.232 |
| *Onychoteuthidae* | 2.31 ± 0 (1) | 8.79 ± 7.55 (16) | 5.7 ± 1.69 (6) | 3.44 ± 0.98 (5) | 4.58 ± 0 (2) | 0 ± 0 (0) | 0.144 |
| *Pyroteuthidae* | 7.84 ± 6.26 (7) | 6.76 ± 3.33 (12) | 8 ± 5.36 (28) | 9.04 ± 5.53 (20) | 6.49 ± 3.8 (12) | 6.08 ± 4.67 (14) | 0.81 |
| *Sepiolidae* | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | not collected at this time of day |
| *Thysanoteuthidae* | 0 ± 0 (0) | 3.68 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | only collected in one depth bin |

Considering depth at night, one taxon was not collected in the nighttime samples (Table 15). Seven were collected in only one depth bin. Octopodidae, Pyroteuthidae, and the group of all paralarvae had significant differences in mean abundance. None of the pairs of bins had significantly different abundances for Octopodidae considering the Holm's adjusted p-values. For Pyroteuthidae, the only significant difference in mean abundance was between the 0-20 m depth bin and the 100-130 m depth bin, of which the shallower bin had higher abundance. For the group of all paralarvae, both the 0-20 m and the 20-40 m depth bins had significantly higher mean abundance than the 80-100 m depth bin.

Although few taxa had significant differences in mean abundance among depth bins, likely due to high variability in abundance among samples and/or few samples, qualitative conclusions about vertical distribution can be made (Figure 5). For example, during the day, Enoploteuthidae has fairly even abundance throughout the near-surface water column (0-130 m). However, at night, the abundance in the top two depth bins (0-20 m, 20-40 m) is higher than in the deeper bins suggesting possible diel vertical migration. Similarly, Pyroteuthidae shows relatively even abundance during the day, but at night the abundance in the surface depth bin (0-20 m) is higher than the other bins and the other bins all have lower abundance than they did during the day. Octopodidae showed the highest daytime abundance in one of the middle depth bins (40-60 m) while the other bins had roughly equal abundance during the day. At night, the surface two depth bins (0-20 m, 20-40 m) had higher abundance than the lower bins and no individuals were collected in the deepest depth bin (100-130 m).

Table 15. Mean abundance and standard deviation of each cephalopod taxon in each depth bin in nighttime samples. The number of individuals of each taxon collected in each depth bin in nighttime samples is listed in parentheses. Significant p-values are in bold. These values were calculated from 607 samples.

| *Taxon* | *0-20 m* | *20-40 m* | *40-60 m* | *60-80 m* | *80-100 m* | *100-130 m* | *p-value* |
|---|---|---|---|---|---|---|---|
| *All Paralarvae* | 28.73 ± 30.24 (293) | 33.74 ± 56.41 (436) | 21.32 ± 31.52 (218) | 11.83 ± 8.33 (80) | 10.12 ± 8.24 (60) | 10.82 ± 7.27 (53) | **0.001** |
| *Ancistrocheiridae* | 2.85 ± 0 (1) | 3.5 ± 1.26 (2) | 3.56 ± 0 (1) | 3.58 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) | 0.849 |
| *Argonautidae* | 4.16 ± 0.43 (3) | 3.29 ± 0.65 (3) | 0 ± 0 (0) | 3.15 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) | 0.193 |
| *Brachioteuthidae* | 9.8 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | only collected in one depth bin |
| *Chiroteuthidae* | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 3.25 ± 0 (1) | 3.1 ± 0 (1) | 6.04 ± 0 (2) | 0.368 |
| *Chtenopterygidae* | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 2.26 ± 0 (1) | only collected in one depth bin |
| *Cranchiidae* | 4.02 ± 0 (1) | 3.4 ± 0.36 (4) | 7.62 ± 0.53 (4) | 3.34 ± 0.63 (2) | 3.37 ± 0.56 (3) | 3.4 ± 1.62 (3) | 0.289 |
| *Cycloteuthidae* | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 2.26 ± 0 (1) | only collected in one depth bin |
| *Enoploteuthidae* | 11.8 ± 16.92 (72) | 18.86 ± 45.76 (123) | 6.93 ± 4.66 (27) | 5.86 ± 5.05 (11) | 3.52 ± 0.84 (3) | 3.72 ± 0 (1) | 0.23 |
| *Histioteuthidae* | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 2.94 ± 0 (1) | only collected in one depth bin |
| *Joubiniteuthidae* | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 2.1 ± 0 (1) | only collected in one depth bin |
| *Loliginidae* | 6.18 ± 2.73 (10) | 9.12 ± 11.21 (10) | 4.89 ± 2.42 (8) | 12.06 ± 6.05 (7) | 2.84 ± 0 (1) | 0 ± 0 (0) | 0.282 |
| *Lycoteuthidae* | 3.68 ± 1.06 (2) | 6.8 ± 0 (2) | 5.18 ± 1.78 (4) | 5.42 ± 3.34 (3) | 3.79 ± 0.98 (2) | 2.25 ± 0 (1) | 0.44 |
| *Mastigoteuthidae* | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 2.78 ± 0 (1) | 0 ± 0 (0) | only collected in one depth bin |
| *Neoteuthidae* | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | not collected at this time of day |
| *Octopodidae* | 10.24 ± 10.29 (56) | 13.51 ± 13.47 (73) | 6.09 ± 2.91 (16) | 5.12 ± 4.55 (10) | 3.64 ± 1.34 (3) | 0 ± 0 (0) | **0.043** |
| *Octopoteuthidae* | 3.17 ± 0 (1) | 7.34 ± 7.81 (8) | 3.56 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) | 2.56 ± 0.65 (2) | 0.361 |
| *Ommastrephidae* | 7.32 ± 4.68 (37) | 7.15 ± 5.67 (49) | 7 ± 4.8 (42) | 3.33 ± 0.56 (7) | 4.35 ± 2.04 (12) | 11.28 ± 10.29 (8) | 0.054 |
| *Onychoteuthidae* | 4.85 ± 2.16 (4) | 3.46 ± 0.1 (3) | 3.21 ± 0.45 (4) | 2.73 ± 0.41 (2) | 0 ± 0 (0) | 3.79 ± 2.1 (5) | 0.519 |
| *Pyroteuthidae* | 11.42 ± 11.15 (17) | 4.74 ± 2.76 (17) | 4.16 ± 1.7 (14) | 4.25 ± 1.71 (14) | 5.89 ± 5.25 (16) | 3.95 ± 2.66 (10) | **0.028** |
| *Sepiolidae* | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 3.82 ± 0 (1) | 0 ± 0 (0) | 11.76 ± 0 (4) | 1 |
| *Thysanoteuthidae* | 3.62 ± 0 (1) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | 0 ± 0 (0) | only collected in one depth bin |

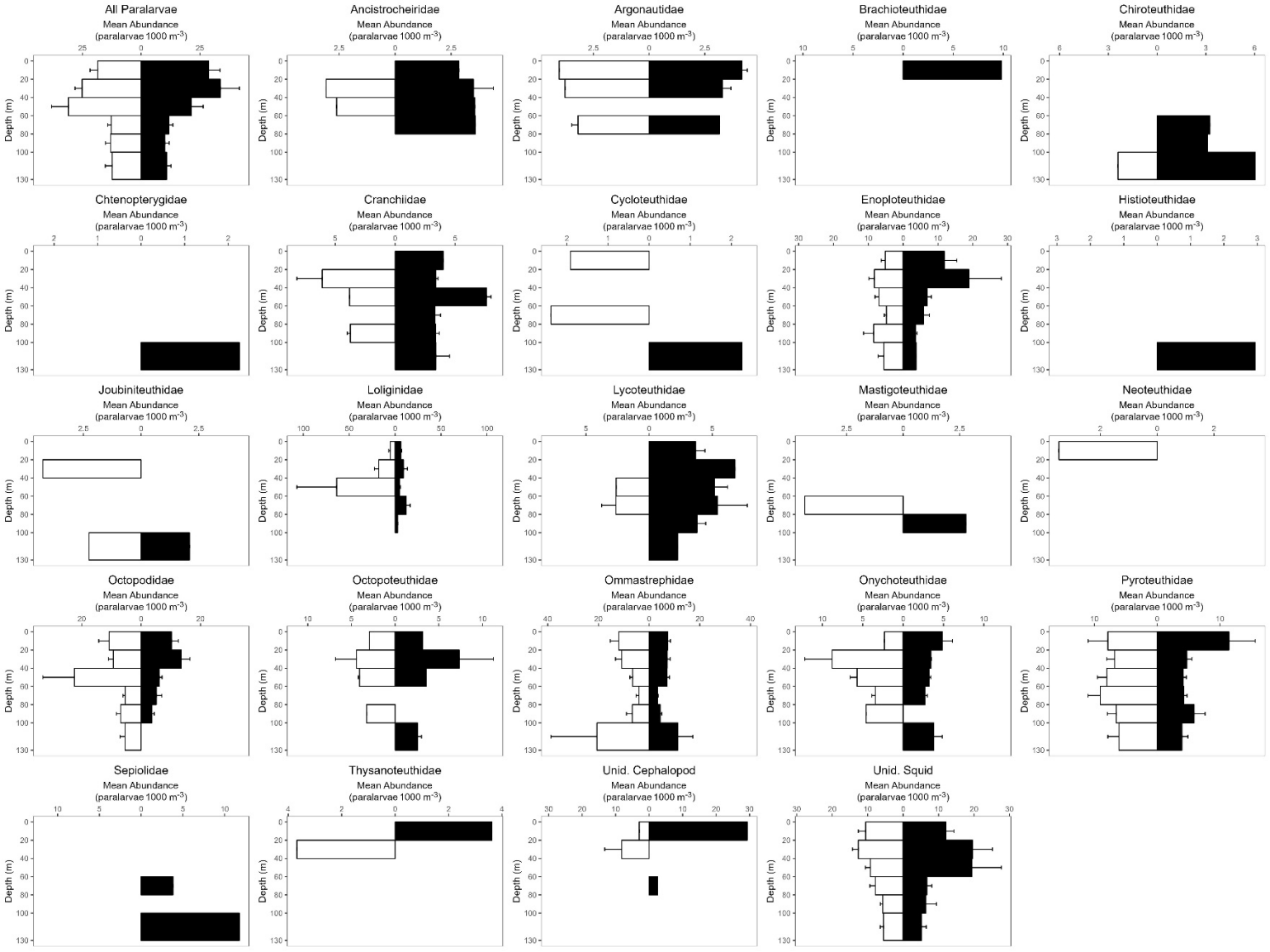


Figure 5. Mean abundance and standard error in each depth bin by time of day for all taxa. The white bars (left) indicate daytime values while the black bars (right) indicate nighttime values. The range of abundance values on the x-axis varies between taxa.

Of the sixteen taxa collected at both times of day, ten, as well as the all paralarvae group, had a deeper abundance-weighted mean depth during the day than at night (Figure 6). This pattern is more evident for the frequently occurring taxa, with eight of the ten most frequently occurring taxa having a deeper abundance-weighted mean depth during the day. The largest differences among the most frequently occurring taxa were for Octopodidae and Argonautidae, for both of which the abundance-weighted mean depth is 17 m deeper in day versus night and Enoploteuthidae, for which the abundance-weighted mean depth is 14 m deeper in day versus night. The two taxa of the most frequently occurring group with deeper abundance-weighted mean depths at night were Ommastrephidae, 4 meters deeper, and Onychoteuthidae, 9 meters deeper.

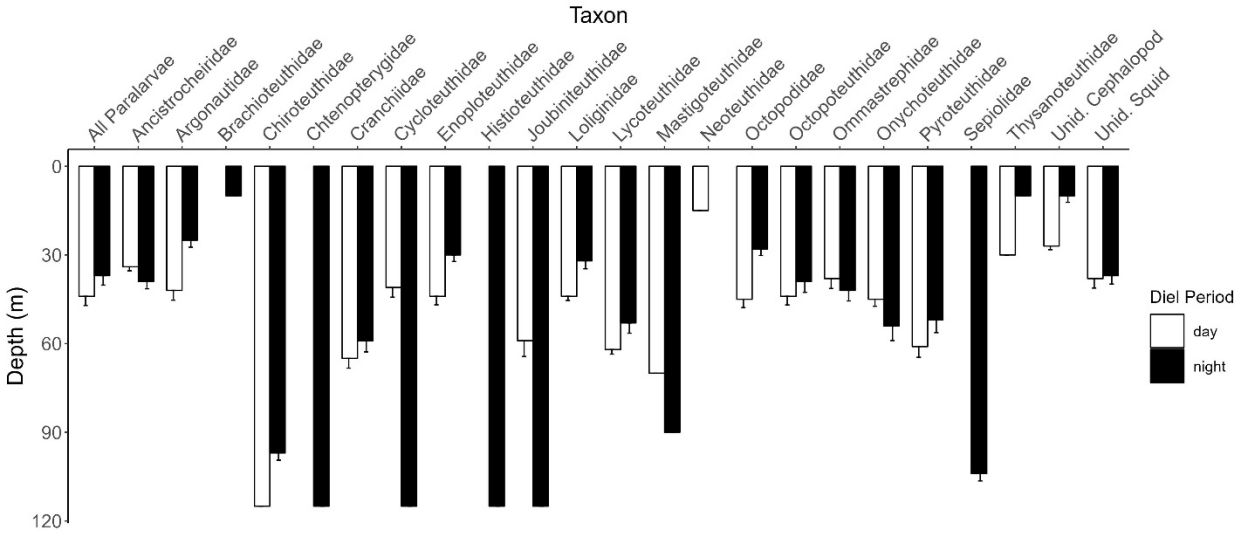


Figure 6. Abundance-weighted mean depth during the day and at night for each taxon.

## 3.3. Abundance Modeling

### *3.3.1. Data Exploration*

The abundance modeling process started with 607 samples that fit within the six depth bins. Of these, 80 samples had missing CTD data and were removed from the dataset, resulting in 527 samples for analysis. Eleven samples had outlying values of salinity and transmissivity

and were removed from the dataset, resulting in 516 samples. Four variables had initial modified VIFs above the cut-off of 4: density, temperature, salinity, and season. After sequential removal of density and temperature, none of the remaining variables had a modified VIF above the cut-off and all were included in the model. Sampling effort varied considerably in time and space (Appendix 2).

### *3.3.2. Enoploteuthidae*

The response variable, count of Enoploteuthidae, had three outlying values (26 paralarvae, 32 paralarvae, and 54 paralarvae) while all other values were less than 10. A high percentage of the samples (76%) contained zero Enoploteuthidae paralarvae. Sites with at least one non-zero count appeared to be more likely to have other samples with non-zero counts, suggesting there may be patterns due to repeated measurements from the same object (site) (Appendix 2). The refined negative binomial model had a positive adjusted $R^2$ value and the dispersion statistic closest to 1 of the three models and was selected for further model fitting. The dispersion statistic was 1.08, indicating that the model was likely not overdispersed.

The best-fitting model produced by backwards selection of variables included volume filtered as a smooth term and year, season, diel period, and depth bin. The dispersion statistic of this model was 1.179, indicating that it is likely not overdispersed. This model explained 21.4% of the deviance in the data while the adjusted $R^2$ was 0.0612. The estimated value of θ was 0.289. Volume filtered showed a bimodal behavior, with the highest expected counts when volume filtered is approximately 225 $m^3$ or above 425 $m^3$, although the confidence interval is very wide on the right side of the plot (Figure 7). Counts did not differ significantly between years. Counts were statistically equal in winter and early fall, but significantly lower in spring. Nighttime counts were significantly higher than daytime counts. Counts in the four shallowest bins are not different, while those in the two deepest bins are significantly lower. Detailed model output can be found in Appendix 2.

Model validation indicates slight heterogeneity of variance along with non-normality in the error terms (Appendix 2). There were no obvious patterns between the residuals and any of the variables in the model, indicating good model fit. Additionally, there were no obvious patterns between the residuals and any of the variables not in the model, indicating that none of the variables needed to be added into the model. The Cook's distances suggest that there are no strongly influential points, but there may be weakly influential observations.

The Moran's I calculated on counts between sites was -0.004 with a p-value of 0.743. Therefore, the null hypothesis of zero spatial autocorrelation cannot be rejected, meaning the count of Enoploteuthidae between different sites is not correlated based on distance. Finally, the Moran's I on counts within sites was 0.468 with a p-value of 0. The null hypothesis of zero spatial autocorrelation is thus rejected. Therefore, the count of Enoploteuthidae in two samples taken at the same site is more likely to be similar than in a sample taken at a different site. This suggests that a random intercept should be added to the model to account for correlation between samples taken at the same site.

### *3.3.3. Octopodidae*

The response variable count of Octopodidae ranged from 0 to 20 with no outliers. A high percentage of the samples (77%) contained zero Octopodidae paralarvae. Sites with at least one non-zero count appeared to be more likely to have other samples with non-zero counts, suggesting there may be patterns due to repeated measurements from the same object (site) (Appendix 2). The refined negative binomial model had a positive adjusted $R^2$ value and the dispersion statistic closest to 1 of the three models and was selected for further model fitting. The dispersion statistic was 1.275, indicating possible slight overdispersion.

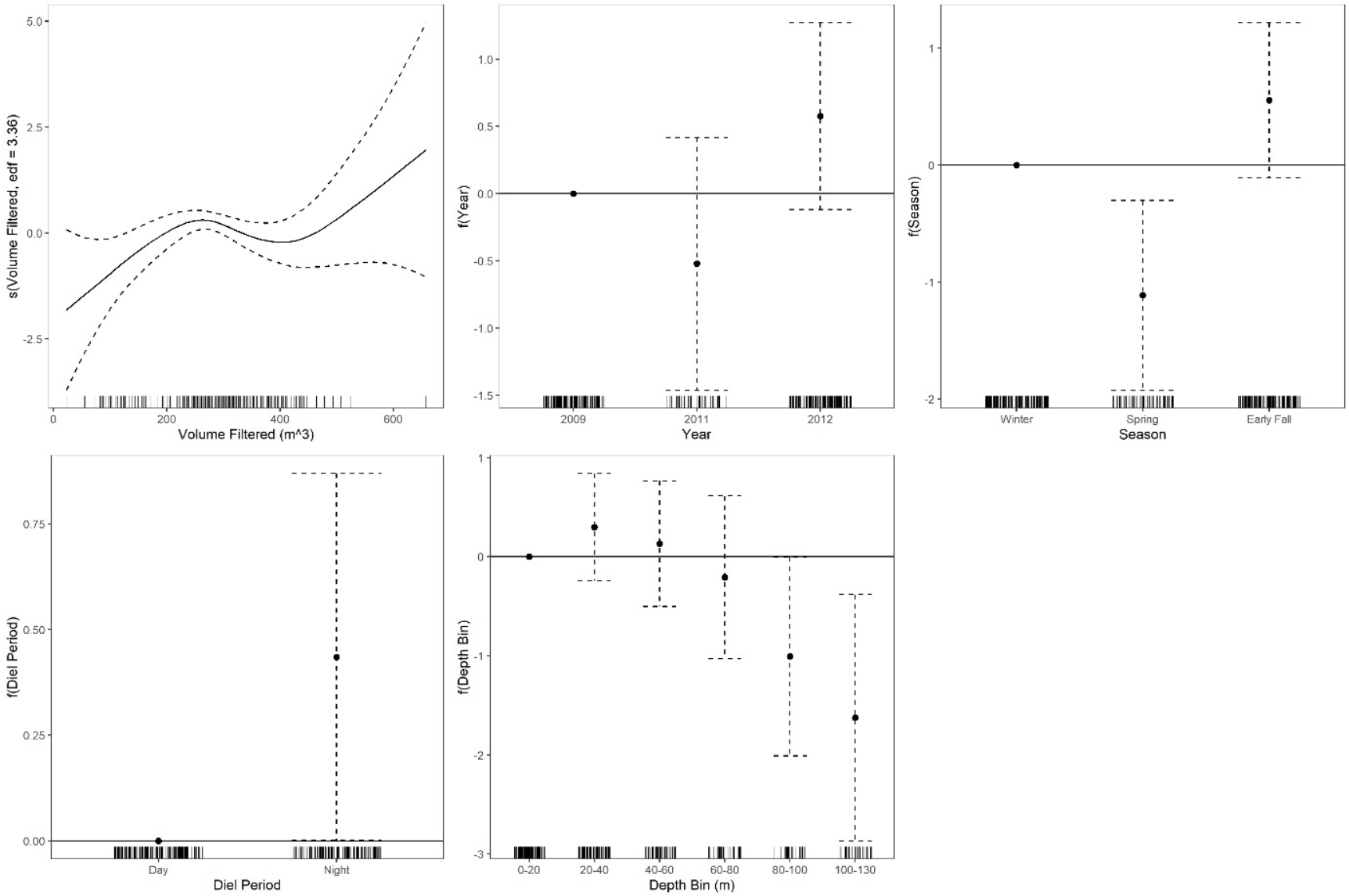


Figure 7. Fitted smooth effect of volume filtered and predictor of each parametric model component of the Enoploteuthidae count model. Rug plot (at the foot of each plot) shows observations of the explanatory variable.

The best-fitting model produced by backwards selection included volume filtered, fluorescence, season, and region, all as parametric terms. As no smoothers were included in the final model, it is essentially a generalized linear model. The dispersion statistic of this model was 1.207, indicating the model may be slightly overdispersed. This model explained 23.8% of the deviance in the data while the adjusted $R^2$ was 0.122. The estimated value of θ was 0.287. Count of Octopodidae increased slightly with increasing volume filtered and increased at a greater rate with increasing fluorescence (Figure 8). Counts are expected to be the same in the NW GOM shelf and WFS, but significantly lower in the other regions. Counts are expected to be the same in winter and spring, but significantly higher in early fall. Detailed model output can be found in Appendix 2.

Model validation indicates slight heterogeneity of variance along with non-normality in the error terms (Appendix 2). There were no obvious patterns between the residuals and any of the variables in the model, indicating good model fit. Additionally, there were no obvious patterns between the residuals and any of the variables not in the model, indicating that none of the variables needed to be added into the model. The Cook's distances suggest that there are no strongly influential points, but there may be weakly influential observations.

The Moran's I calculated on counts between sites was 0.0189 with a p-value of 0.0282. Therefore, the null hypothesis of zero spatial autocorrelation is rejected, meaning the count of Octopodidae between different sites is slightly but significantly correlated based on distance. Sites close together are more likely to have similar counts of Octopodidae than sites farther away. This suggests that a spatial autocorrelation structure should be added to the model. Finally, the Moran's I calculated on counts within sites was 0.1748 with a p-value of 1.9645e-09. The null hypothesis of zero spatial autocorrelation is therefore rejected. The count of Octopodidae in two samples taken at the same site is thus more likely to be similar than in a sample taken at a different site. This suggests that a random intercept should be added to the model to account for correlation between samples taken at the same site.

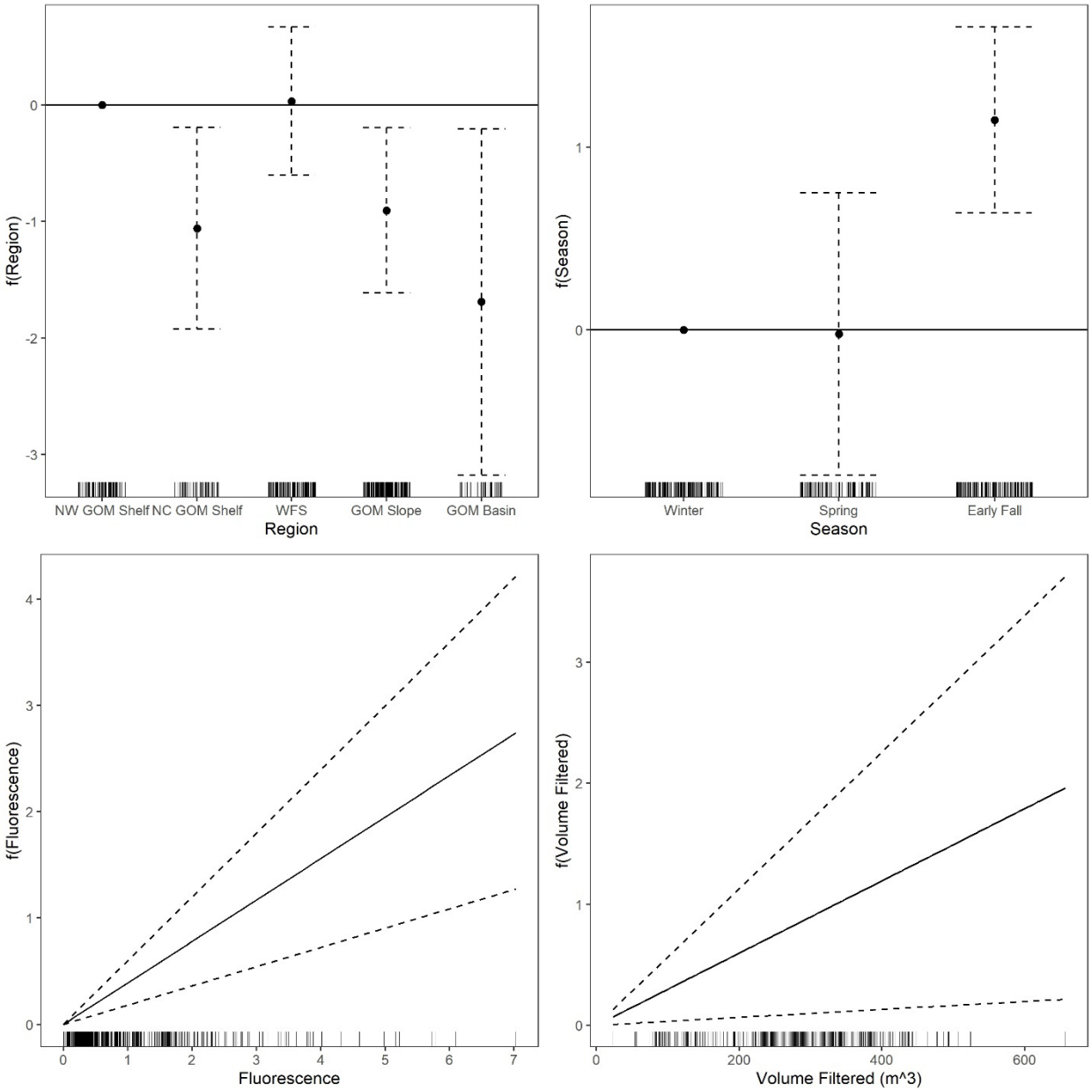


Figure 8. Predictor of each parametric model component of the Octopodidae count model. Rug plot (at the foot of each plot) shows observations of the explanatory variable.

### *3.3.4. Loliginidae*

The response variable, count of Loliginidae, had one outlying value of 47 paralarvae, while all other values were less than 10 paralarvae. A very high percentage of the samples (93%) contained zero Loliginidae paralarvae. Sites with at least one non-zero count appeared to be more likely to have other samples with non-zero counts, suggesting there may be patterns

due to repeated measurements from the same object (site) (Appendix 2). The refined quasi-Poisson model had a positive adjusted $R^2$ value and the dispersion statistic closest to 1 of the three models and was selected for further model fitting. The dispersion statistic was 2.33, indicating that the model was still overdispersed.

The best-fitting model produced by backwards selection included transmissivity and fluorescence as smooth terms, volume filtered as a parametric term, and year. The dispersion statistic of this model was 1.780, indicating that the model was still overdispersed. This model explained 66% of the deviance in the data while the adjusted $R^2$ was 0.75. Transmissivity showed a highly variable pattern, with the highest expected counts when transmissivity is approximately 91% or 97% (indicating clear water), with a decrease between those values. The confidence interval is very wide at values below approximately 87%, indicating high uncertainty in that range (Figure 9). Fluorescence showed an almost linear positive effect, with the highest expected counts when fluorescence is between approximately 3.5 and 6. Counts are expected to decrease slightly with increasing volume filtered. Counts are expected to be significantly higher in 2011 or 2012 than in 2009. As the model is still overdispersed, the quasi-Poisson distribution is likely not the most appropriate distribution for these data. Therefore, the results of the final model should be interpreted with caution. Detailed model output can be found in Appendix 2.

Model validation indicates serious heterogeneity of variance along with non-normality in the error terms (Appendix 2). There were no obvious patterns between the residuals and any of the variables in the model, indicating good model fit. There were no obvious patterns between the residuals and any of the variables not in the model, except for diel period. The F-test for this variable was significant, indicating that there is a relationship between the residuals and diel period, which suggests that it should have been included in the model. The Cook's distances suggest that there are two strongly influential points, which have a noticeable effect on the

model produced on the set of samples containing them. Additionally, there may be weakly influential observations.

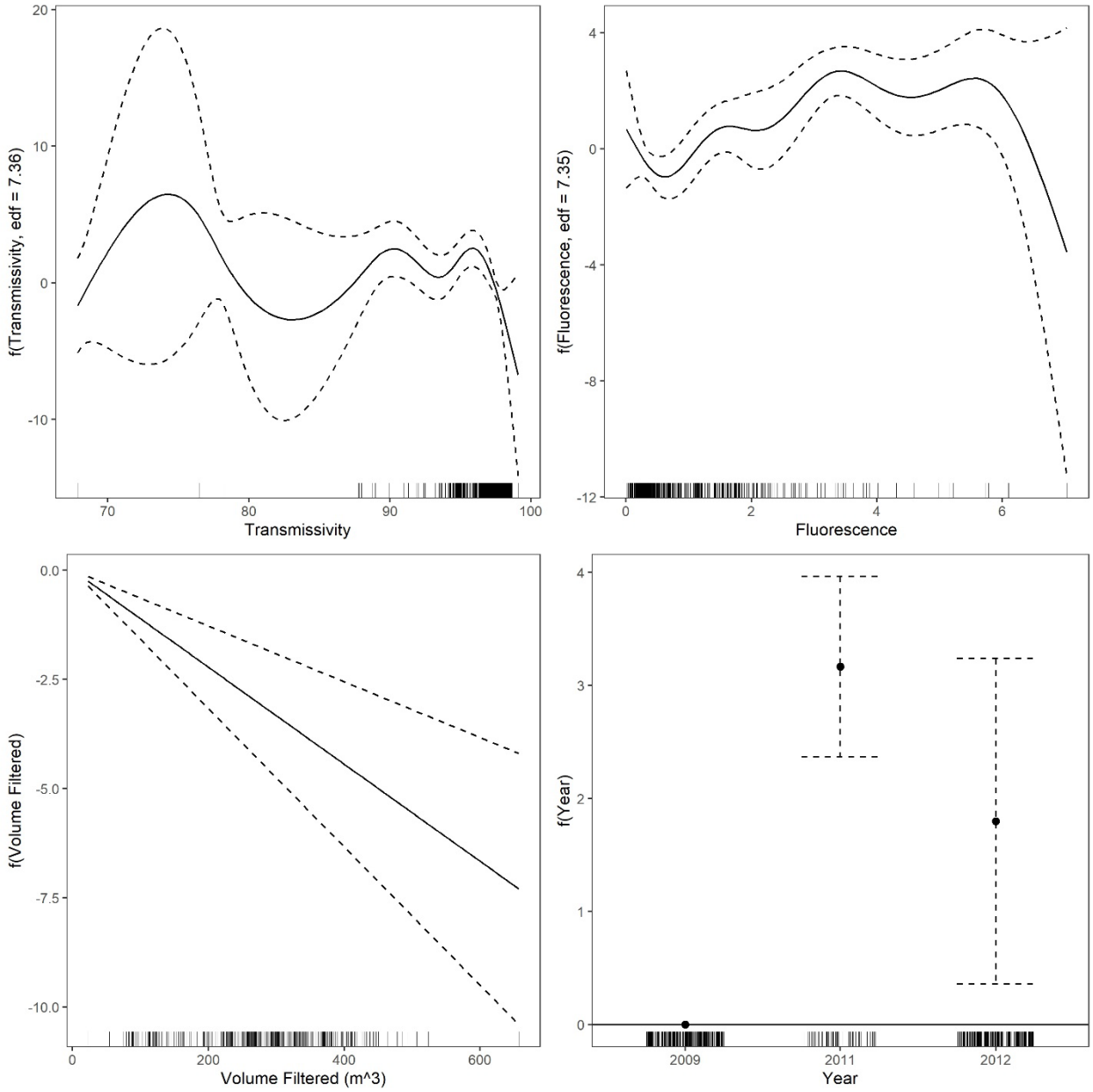


Figure 9. Fitted smooth effects of transmissivity and fluorescence and predictors of each parametric model component of the Loliginidae count model.

The Moran's I calculated on counts between sites was 0.00290 with a p-value of 0.308. Therefore, the null hypothesis of zero spatial autocorrelation cannot be rejected, meaning the count of Loliginidae is not correlated based on distance. Sites close together are no more likely to have similar counts of Loliginidae than sites farther away. Finally, the Moran's I calculated on counts within sites was 0.0676 with a p-value of 1.99e-06. This means that the null hypothesis of zero spatial autocorrelation is rejected. Therefore, the count of Loliginidae in two samples taken at the same site is more likely to be similar than for a sample taken at a different site. This suggests that a random intercept should be added to the model to account for correlation between samples taken at the same site.

### 3.4. Community Structure

A total of 361 samples containing at least one identified paralarva were used to test year, season, diel period, and region effects, while 343 samples were used to test for effects of depth. The PCOA ordination diagram does not show any obvious groupings of samples, although there does appear to be some slight separation between clusters of samples with similar taxa composition and abundance. (Figure 10). The first principal coordinate axis explains only 7.72% of the variability in the composition and abundance of the samples and the second PCO-Axis explains only 6.84% of the variability. These two values are very close, so units on the two axes are approximately equivalent.

The original species weighted biplot vectors figure shows only five taxa with enough influence to be visibly separate from the cluster of taxa in the middle of the diagram (Figure 10). These are the most commonly collected taxa: Enoploteuthidae, Loliginidae, Octopodidae, Ommastrephidae, and Pyroteuthidae. Of these, Octopodidae is the single most important taxon driving the observed patterns among samples because its vector has the longest component along the (marginally) most important axis (axis-I). It also has the third longest component along the second-most important axis. Enoploteuthidae and Ommastrephidae have about the same

length along axis-I, but Enoploteuthidae is longer along axis-II so it is the second most influential taxon in driving the patterns. Ommastrephidae is slightly more influential than Loliginidae and Pyroteuthidae is similar or even slightly more influential than them.

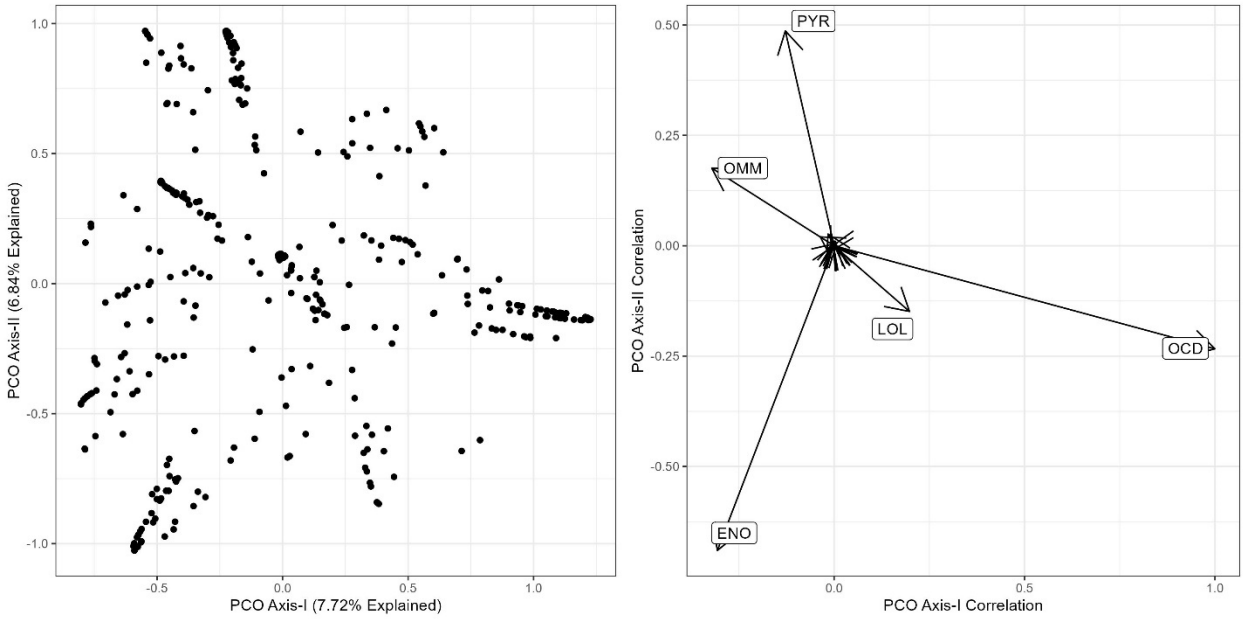


Figure 10. PCOA ordination diagram and species weighted biplot vector plot (non-transformed abundance data). ENO – Enoploteuthidae, LOL – Loliginidae, OCD – Octopodidae, OMM – Ommastrephidae, PYR – Pyroteuthidae.

The species weighted biplot vector plot with the vectors fourth-root transformed shows which taxa are likely to be collected in the same samples (Figure 11). Loliginidae and Octopodidae are likely to be found together and far less likely to be found with any of the other taxa. Pyroteuthidae and Ommastrephidae are more likely to be found in the same samples than other taxa. Onychoteuthidae, Cranchiidae, Sepiolidae, Chiroteuthidae, Octopoteuthidae, Cycloteuthidae, Histioteuthidae, Chtenopterygidae, and Joubiniteuthidae are more likely to be found with Ommastrephidae and Pyroteuthidae than other taxa, but because most of those taxa have very short vectors, any inferences made about them are not robust. Based on the direction of its vector, Enoploteuthidae may be found alongside Octopodidae and Loliginidae or Ommastrephidae, Pyroteuthidae, and the cluster of taxa in the direction of Ommastrephidae

and Pyroteuthidae. Lycoteuthidae and Thysanoteuthidae are more likely to be found alongside Enoploteuthidae, but with very short vectors, this inference is not robust. The taxa Argonautidae, Ancistrocheiridae, Brachioteuthidae, Mastigoteuthidae, and Neoteuthidae are less likely to be found alongside Enoploteuthidae, Lycoteuthidae, and Thysanoteuthidae as their vectors point the opposite direction. However, they all have very short vectors so these inferences are not robust.

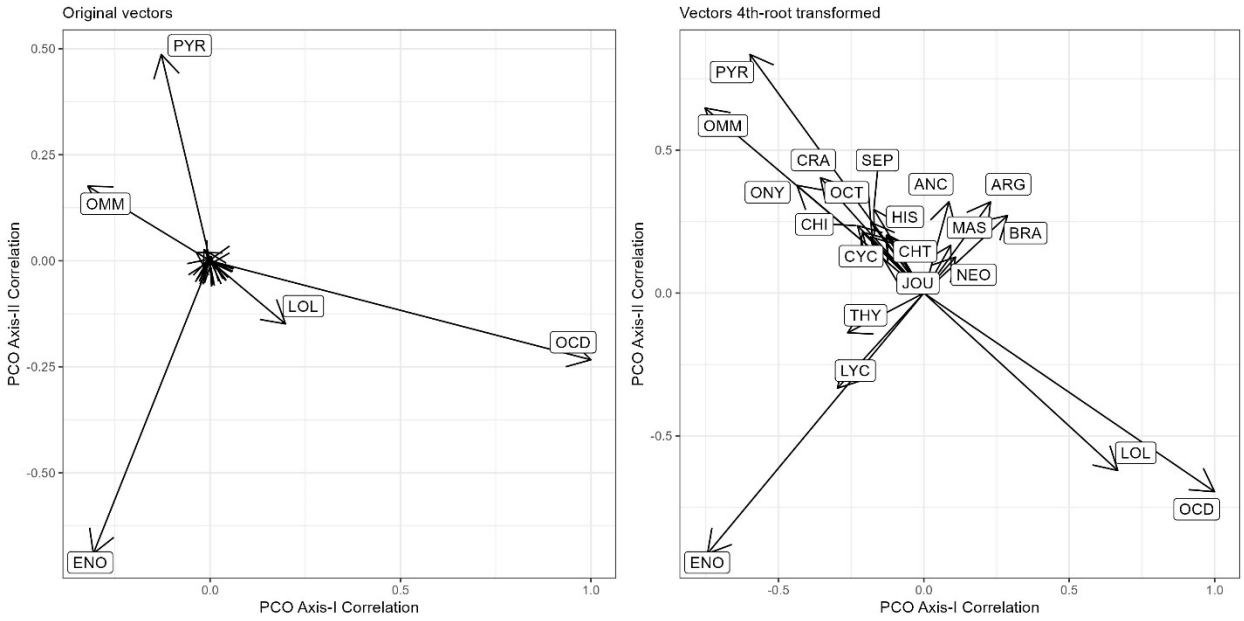


Figure 11. Species weighted biplot vector plots from non-transformed abundance data, with vectors non-transformed, left, and vectors fourth-root transformed, right. ANC – Ancistrocheiridae, ARG – Argonautidae, BRA – Brachioteuthidae, CHI – Chiroteuthidae, CHT – Chtenopterygidae, CRA – Cranchiidae, CYC – Cycloteuthidae, ENO – Enoploteuthidae, HIS – Histioteuthidae, JOU – Joubiniteuthidae, LOL – Loliginidae, LYC – Lycoteuthidae, MAS – Mastigoteuthidae, NEO – Neoteuthidae, OCD – Octopodidae, OCT – Octopoteuthidae, OMM – Ommastrephidae, ONY – Onychoteuthidae, PYR – Pyroteuthidae, SEP – Sepiolidae, THY – Thysanoteuthidae.

For year, the PCOA ellipses for 2009 and 2012 are more similar than that for 2011 and their centroids are closer together than the 2011 centroid is to either of them (Figure 12). This suggests that these years have more similar paralarval communities. The results of the PERMANOVAs suggest that the variability (dispersion) of the family composition and

abundance (beta-diversity) is different between 2011 and the other two years, but doesn't differ between 2009 and 2012 (Table 16). Additionally, the different years could have different variabilities in the beta-diversity and different average beta-diversities, or just different variabilities (dispersions), since the assumption of homogeneity of multivariate dispersions was not met.

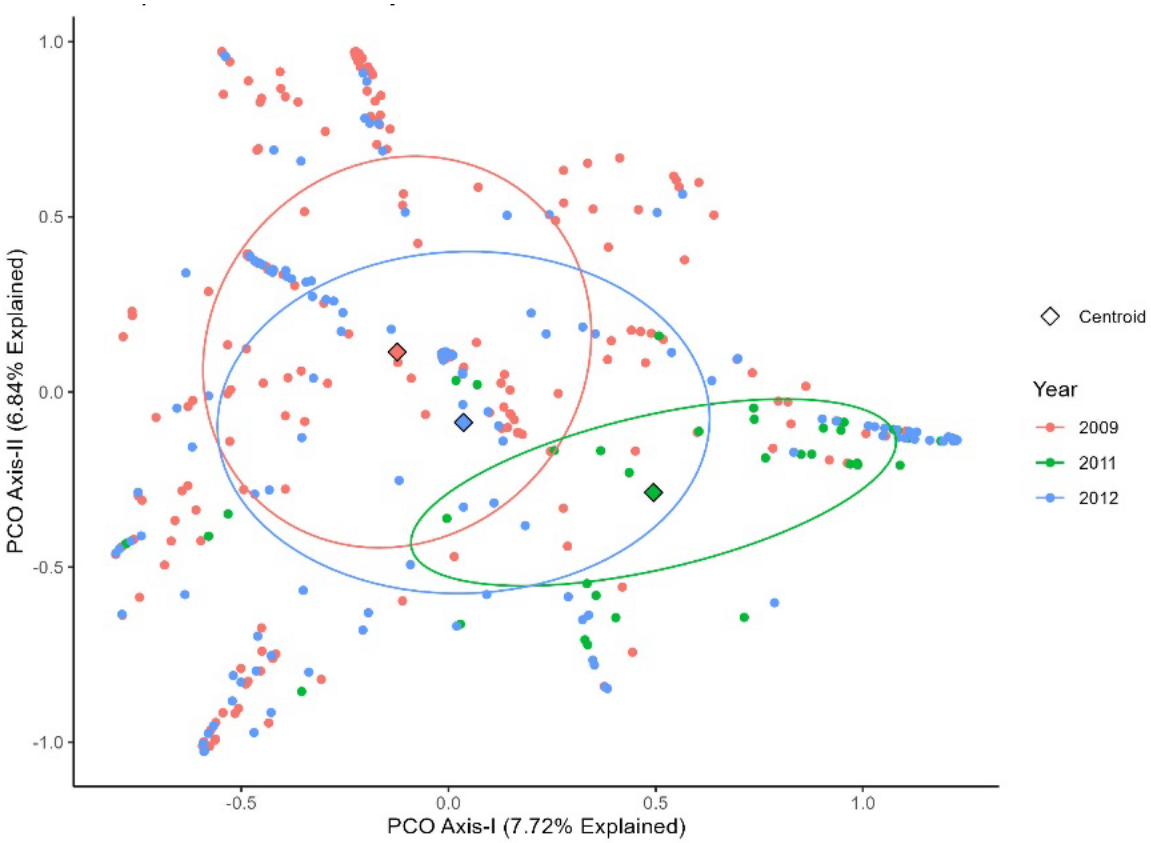


Figure 12. PCOA ordination diagram with samples coded by year collected.

Table 16. Tests for significant differences in the variability of beta-diversity and the average beta-diversity of the paralarval community between years. Significant p-values are bolded. Pairwise p-values are adjusted using Holm's method.

| | | | *2009 vs. 2011* | *2009 vs. 2012* | *2011 vs. 2012* |
|---|---|---|---|---|---|
| | F-statistic | p-value | Adj. p-value | Adj. p-value | Adj. p-value |
| *Variability (dispersion)* | 15.582 | **0.0001** | **3.955e-06** | 0.1852 | **5.647e-08** |
| *Average beta-diversity (location)* | 13.011 | **0.0001** | **0.0003** | **0.0003** | **0.0003** |

The CAP ordination does not show much separation between the groups, suggesting the communities do not differ greatly (Figure 13). Four taxa had significant indicator power values for the year grouping factor, but all had indicator power values less than 0.5, suggesting that they occurred in multiple years and/or not in a majority of samples from a given year (Table A3.1). Pyroteuthidae predominantly characterizes 2009 samples, Loliginidae and Octopodidae are indicators for 2011 samples, and Octopoteuthidae characterizes 2012 samples.

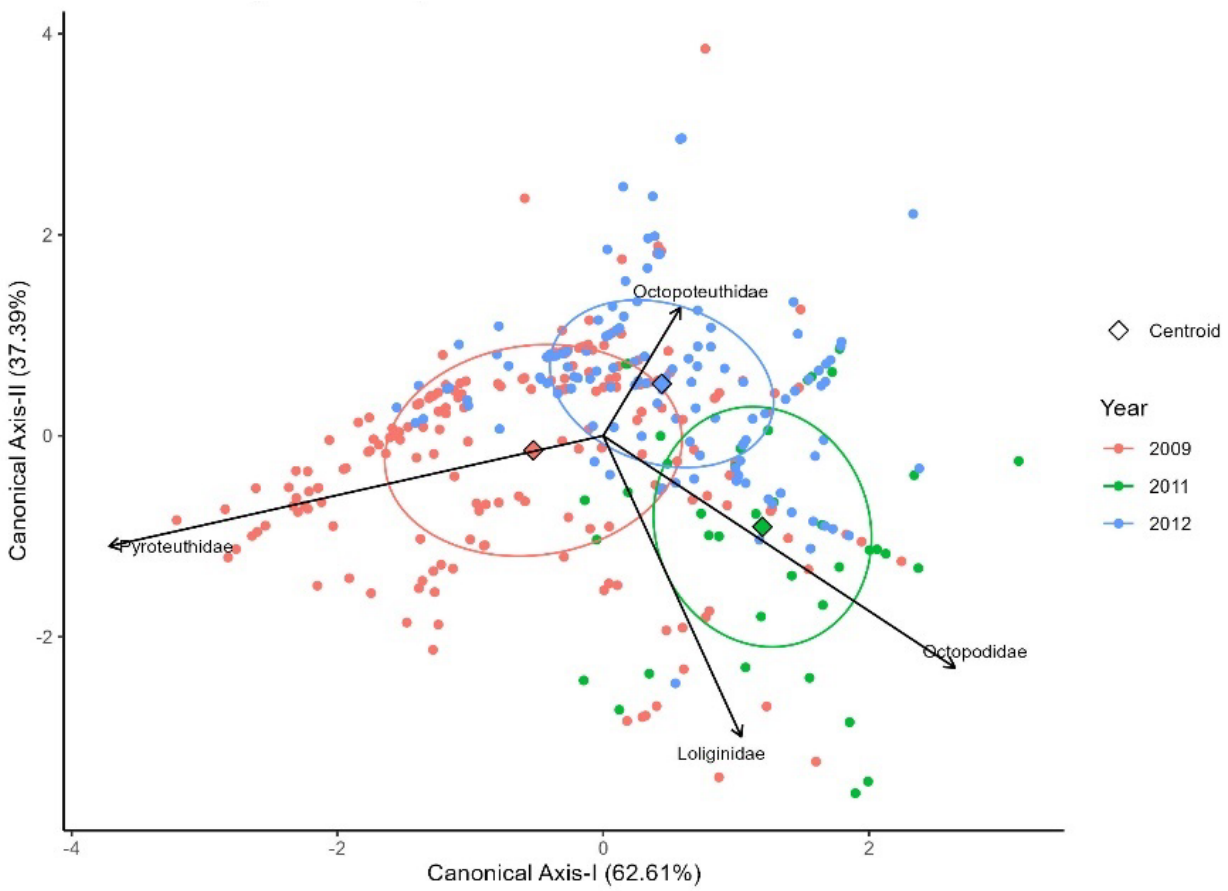


Figure 13. Canonical analysis of principal coordinates ordination for year.

For season, the winter and spring ellipses have fairly similar shapes and sizes, while the early fall ellipse differs from the other two, suggesting that the variability of the beta-diversity is different between early fall and the other seasons, but not between winter and spring; however this is not supported by the PERMANOVA on dispersion (Table 17). All three centroids seem

equidistant, suggesting that all seasons have different paralarval communities (Figure 14). Additionally, the different seasons could have different variabilities (dispersions) and different average beta-diversities, or just different variability in beta-diversity, since the assumption of homogeneity of multivariate dispersions was not met.

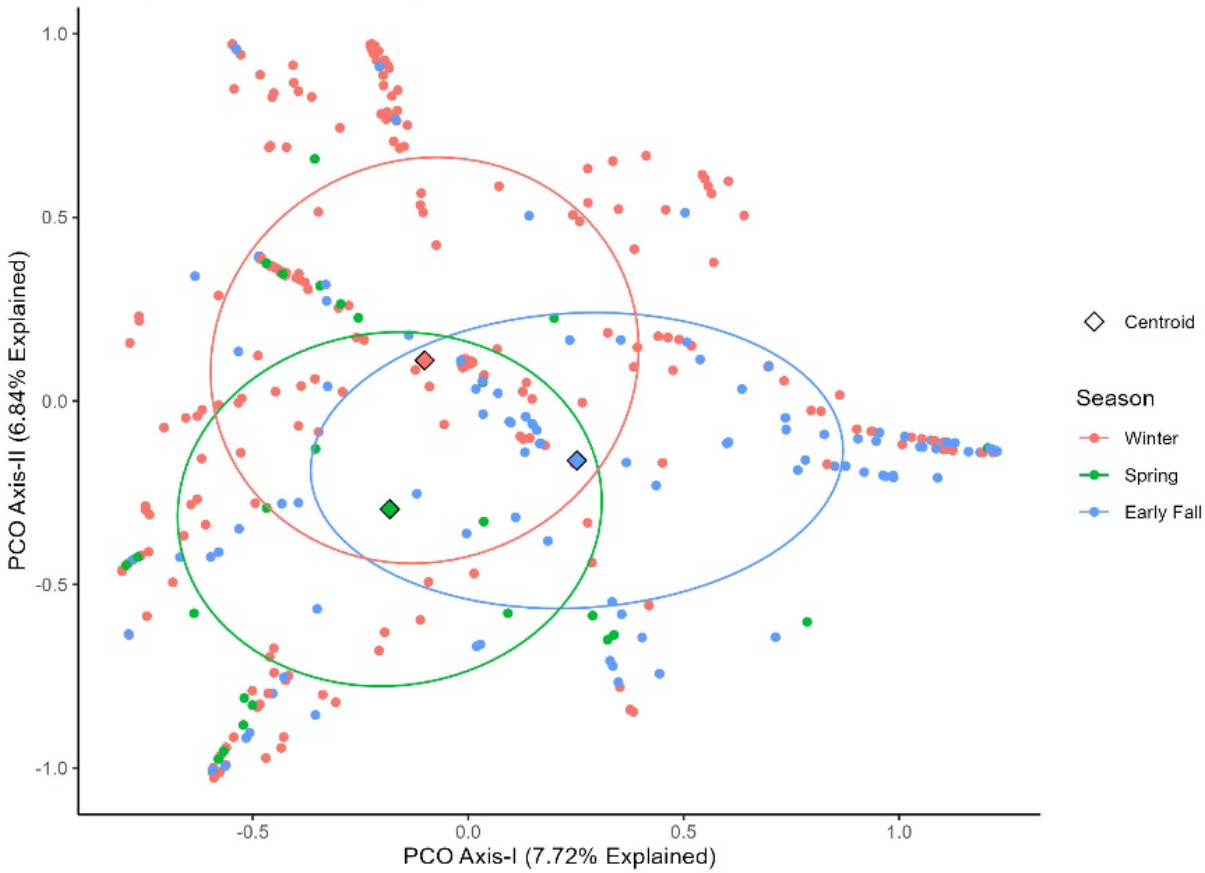


Figure 14. PCOA ordination diagram with samples coded by season collected.

Table 17. Tests of significant differences in the variability of beta-diversity and average beta-diversity of the paralarval community between seasons. Significant p-values are bolded. Pairwise p-values are adjusted using Holm's method.

| | | | *Winter vs. Spring* | *Winter vs. Early Fall* | *Spring vs. Early Fall* |
|---|---|---|---|---|---|
| | F-statistic | p-value | Adj. p-value | Adj. p-value | Adj. p-value |
| *Variability (dispersion)* | 3.3005 | **0.0385** | **0.0128** | 0.3079 | 0.0773 |
| *Average beta-diversity (location)* | 11.725 | **0.0001** | **0.0006** | **0.0003** | **0.0006** |

The CAP ordination does not show much separation between the groups, suggesting the communities do not differ greatly (Figure 15). Six taxa had significant indicator power values for the season grouping factor, but all had indicator power values less than 0.4, suggesting that they occurred in multiple seasons and/or not in a majority of samples from a given season (Table A3.2). Pyroteuthidae is characteristic of winter samples, Ommastrephidae, Enoploteuthidae, and Ancistrocheiridae are representative of spring samples, and Loliginidae and Octopodidae are characteristic of early fall samples.

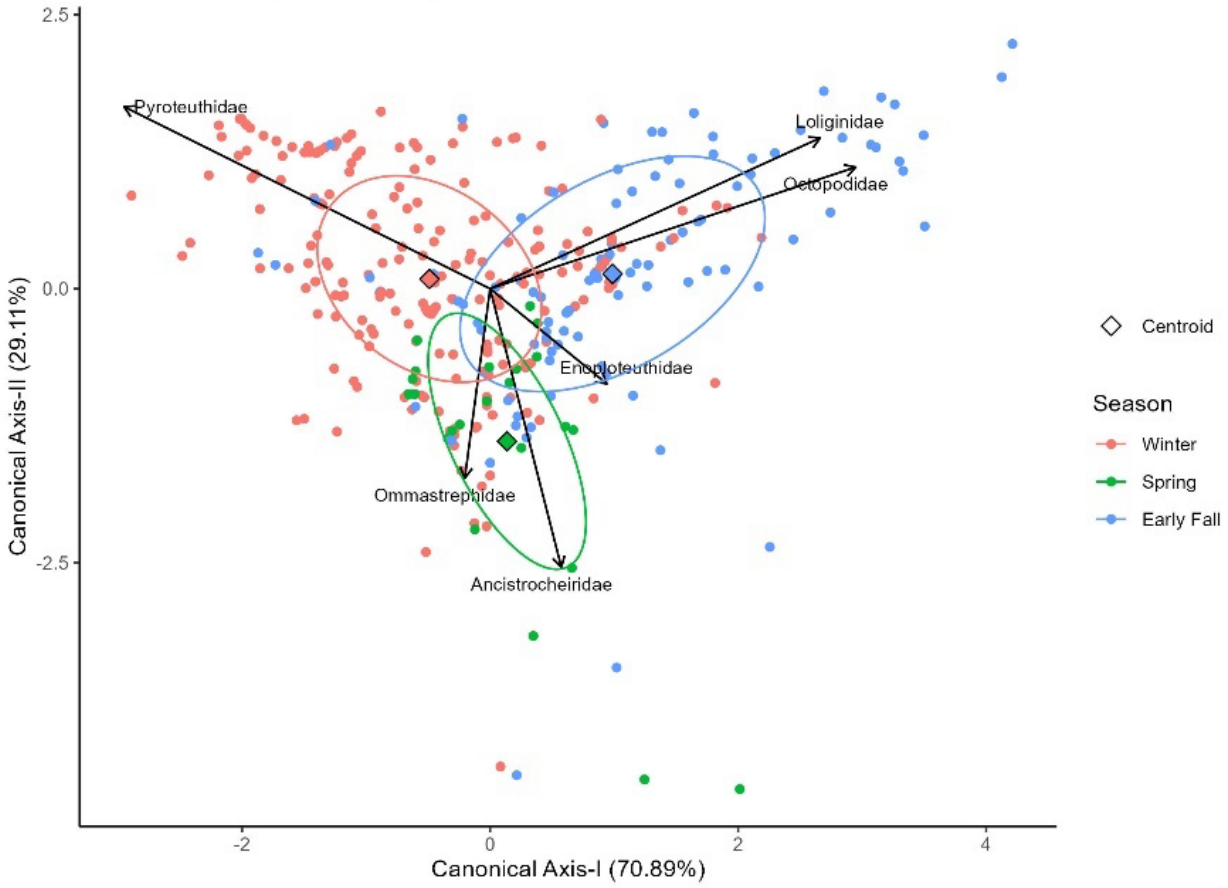


Figure 15. Canonical analysis of principal coordinates ordination for season.

For diel period, the shapes and sizes of the ellipses are similar and the centroids are close, but not overlapping, which suggests that their variability (dispersion) is the same, but they may have significantly different average beta-diversity (locations) (Figure 16). This is supported

by the results of the PERMANOVAs, which found no difference in the variability of the beta-diversity but a significant difference in the average beta-diversity (Table 18). Therefore, the composition and abundance of the paralarval community is different at different times of day.

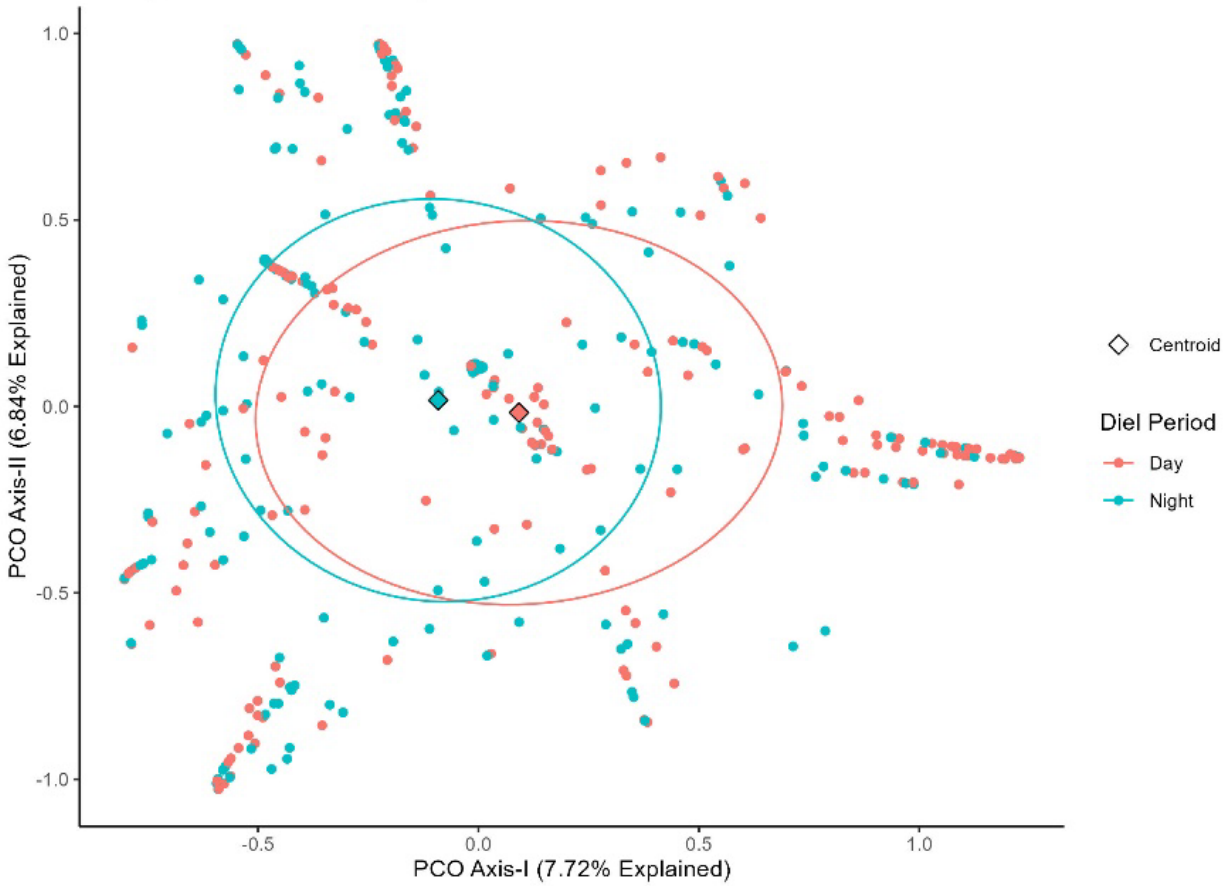


Figure 16. PCOA ordination diagram with samples coded by time of day collected.

Table 18. Tests of significant differences in the variability of beta-diversity and average beta-diversity of the paralarval community between diel periods. Significant p-values are bolded.

| | F-statistic | p-value |
|---|---|---|
| *Variability (dispersion)* | 1.103 | 0.2922 |
| *Average beta-diversity (location)* | 3.085 | **0.0001** |

The CAP ordination shows slight separation between the groups, suggesting differences between the daytime and nighttime communities are minor (Figure 17). Only one taxon, Octopodidae, had a significant indicator power value for the diel period grouping factor (indVal = 0.2577), but the low value suggests that it was found in samples at both times of day and/or not found in a majority of samples from the specified time of day (Table A3.3). It was characteristic of daytime samples.

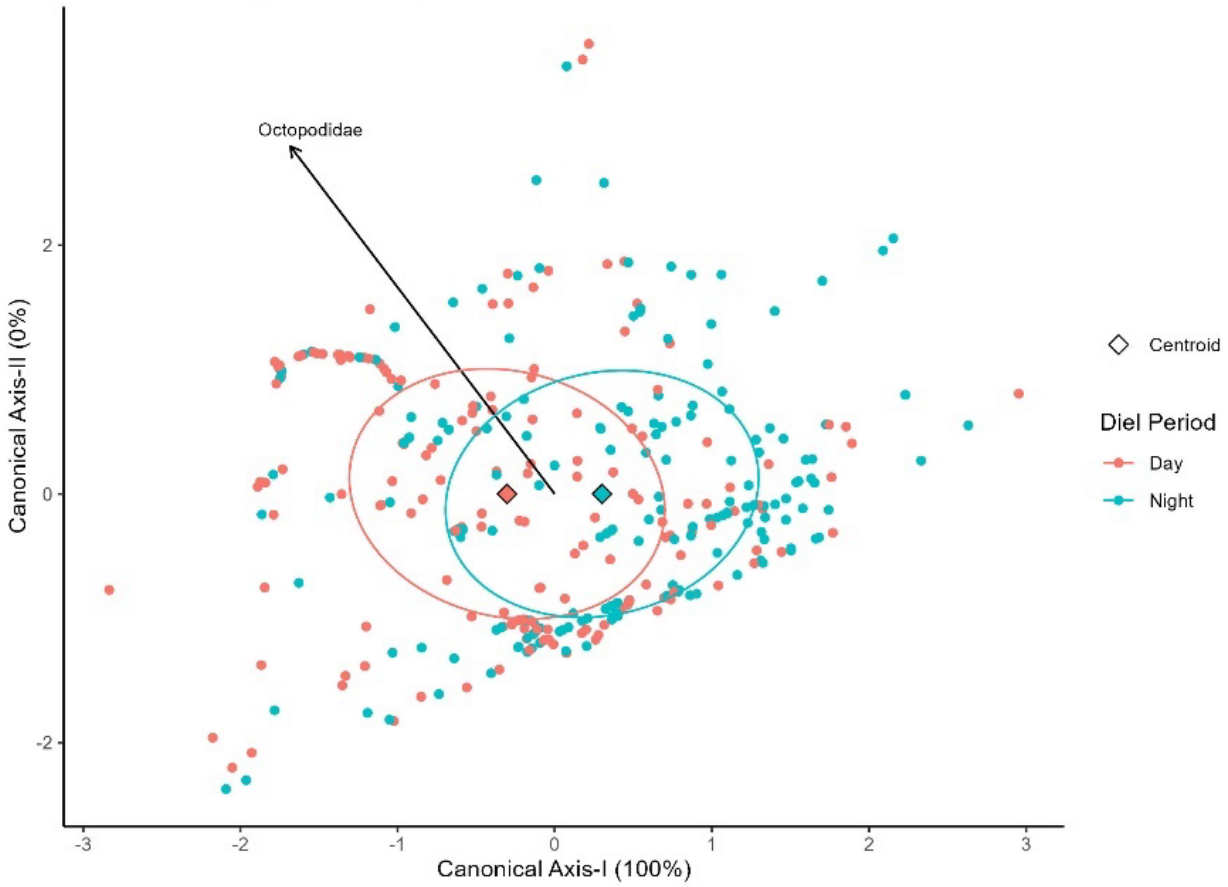


Figure 17. Canonical analysis of principal coordinates ordination for diel period.

The ellipses and centroids for region appear to split into two groups based on the shape and size of the ellipses and location of the centroids: GOM slope and GOM basin vs. NW GOM shelf, North-central GOM shelf, and West Florida Shelf (Figure 18). The results of the PERMANOVA on dispersion indicate that the variability of the beta-diversity is the same in

different regions (Table 19). The results of the PERMANOVA on location (average beta-diversity) support the conclusions from the PCOA that different assemblages of paralarvae are present in different regions. Specifically, the average beta-diversity of the paralarval community is significantly different between the GOM slope and all shelf regions, but not between the GOM slope and GOM basin (Table 20). Additionally, the average beta-diversity differs between the GOM basin and all shelf regions. The average beta-diversity also differed between the NC GOM shelf and the WFS, but not between the NW GOM shelf and NC GOM shelf or the NW GOM shelf and the WFS.

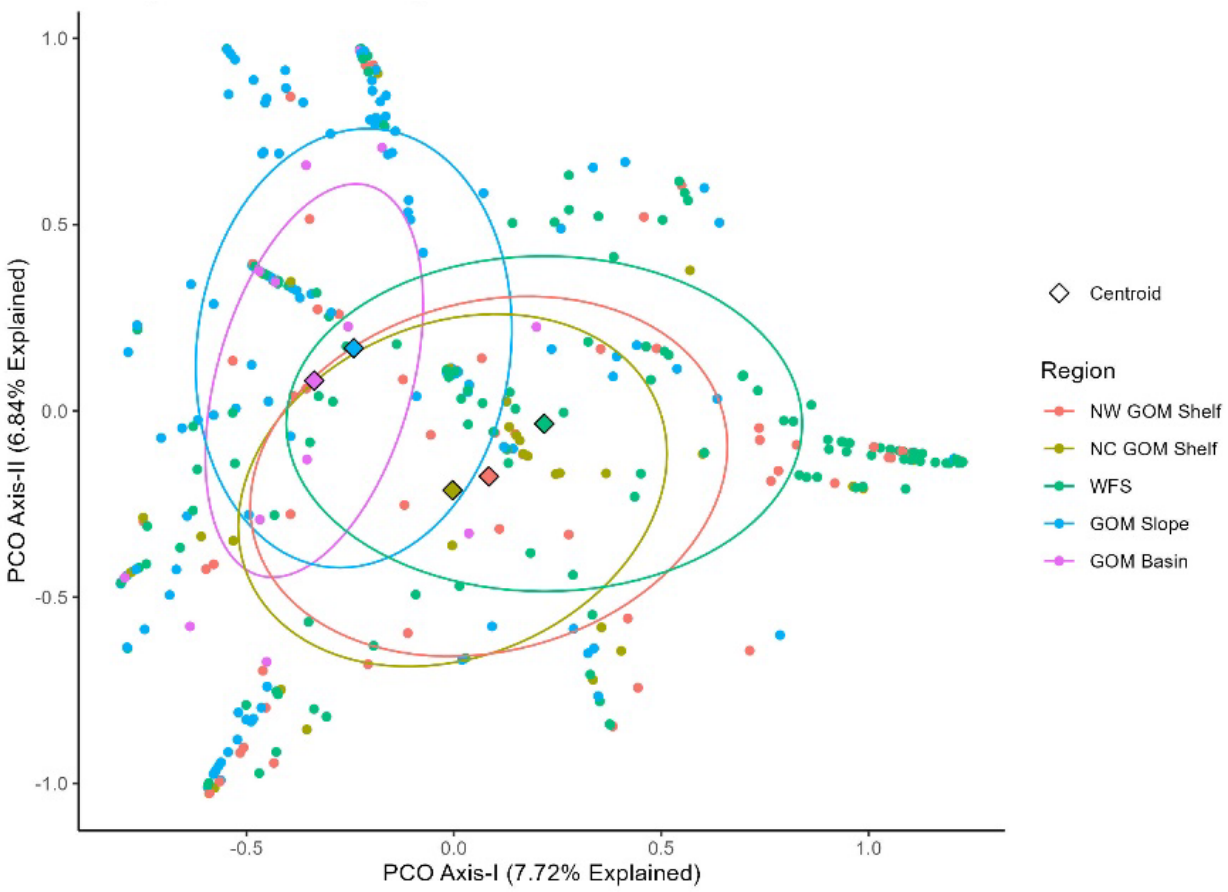


Figure 18. PCOA ordination diagram with samples coded by the region in which they were collected.

Table 19. Tests of significant differences in the variability of beta-diversity (dispersion) and average beta-diversity (location) of the paralarval community between regions. Significant p-values are bolded.

| | F-statistic | p-value |
|---|---|---|
| *Variability (dispersion)* | 0.9398 | 0.4386 |
| *Average beta-diversity (location)* | 7.0781 | **0.0001** |

Table 20. Tests of significant differences in the average beta-diversity (location) of the paralarval community between pairs of regions. Significant p-values are bolded. Pairwise p-values are adjusted using Holm's method.

| | *NW GOM Shelf* | *NC GOM Shelf* | *WFS* | *GOM Slope* |
|---|---|---|---|---|
| *NC GOM Shelf* | 0.2764 | | | |
| *WFS* | 0.1964 | **0.0115** | | |
| *GOM Slope* | **0.0010** | **0.0010** | **0.0010** | |
| *GOM Basin* | **0.0116** | **0.0084** | 0.0084 | 0.1554 |

The CAP ordination shows some separation between GOM slope and GOM basin and the shelf regions, which reflects the results of the PERMANOVA on average beta-diversity, but it is not very distinct (Figure 19). Four taxa had significant indicator power values for the region grouping factor, but all had indicator power values less than 0.4, suggesting that they occurred in multiple regions and/or not in a majority of samples from a given region (Table A3.4). Pyroteuthidae was an indicator of GOM slope samples, Ommastrephidae and Ancistrocheiridae characterized GOM basin samples, and Loliginidae characterized NC GOM shelf samples.

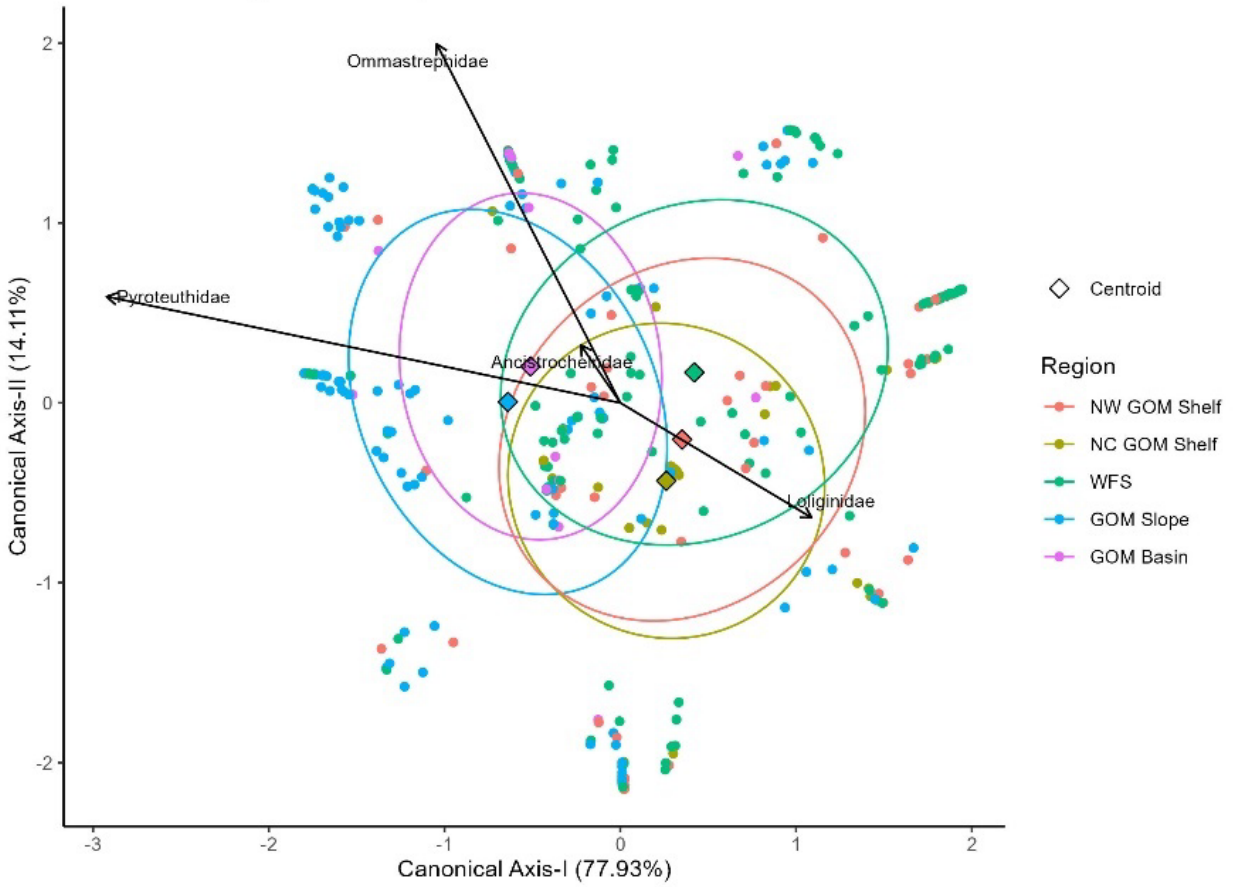

Figure 19. Canonical analysis of principal coordinates ordination for region.

Finally, the ordination for depth bin suggests that the beta-diversity of pairs of depth bins are similar: 0-20 m and 20-40 m, 40-60 m and 60-80 m, and 80-100 m and 100-130 m (Figure 20). The 0-20 m and 20-40 m bins are most different from the 80-100 m and 100-130 m bins while the middle depth bins are in-between the other two groups. The shape of the 80-100 m and 100-130 m ellipses appear different than the others, which suggests they may have different variability in beta-diversity (dispersions). However, this was not supported by the PERMANOVA on dispersion (Table 21). The PERMANOVA on average beta-diversity (locations) indicated that a different assemblage of paralarvae is present in different depth bins. The two shallowest depth bins (0-20 m and 20-40 m) have significantly different average beta-diversity compared to the three deepest bins, but are not different from each other or the third shallowest depth bin (40-60 m) (Table 22). The 40-60 m depth bin only had significantly different average

beta-diversity relative to the deepest depth bin (100-130 m). The three deepest depth bins did not have significantly different average beta-diversities between each other.

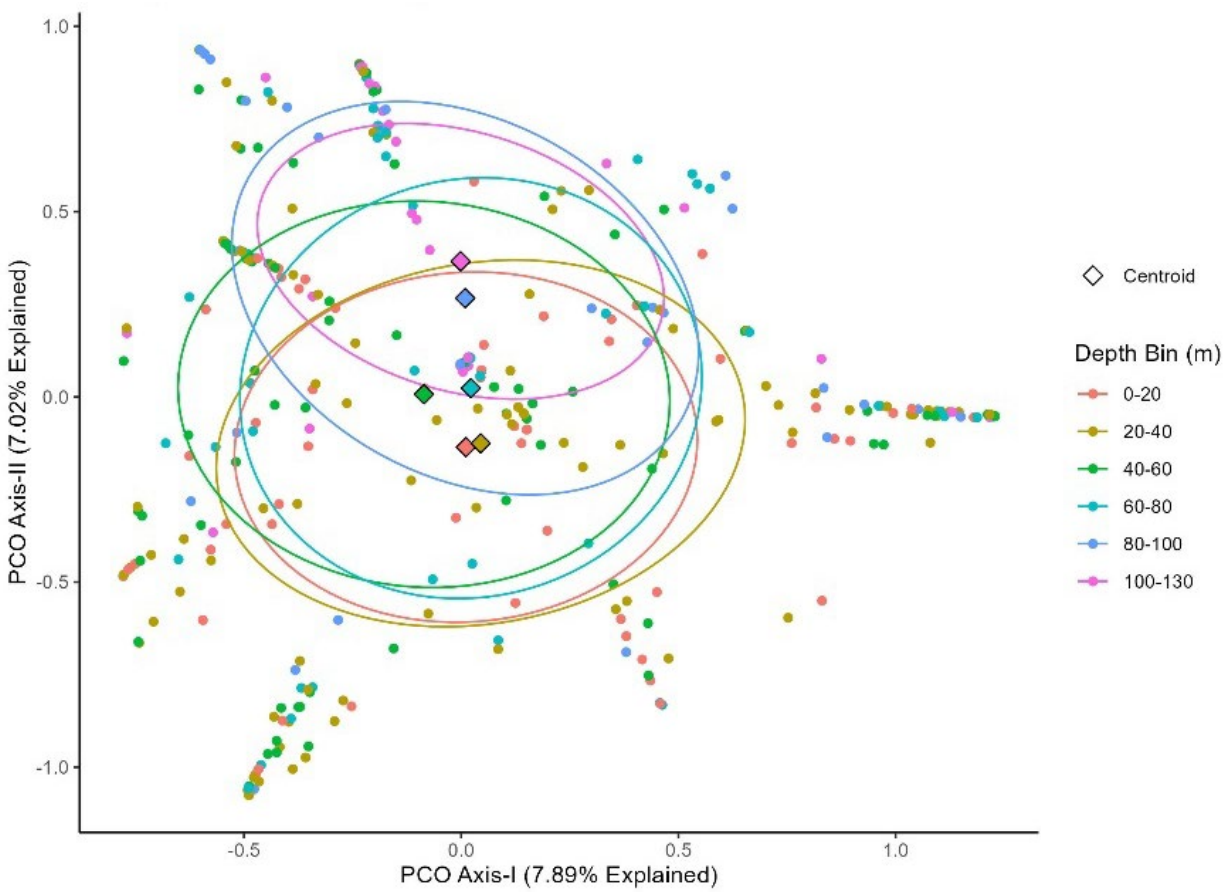


Figure 20. PCOA ordination diagram with samples coded by the depth bin in which they were collected.

Table 21. Tests of significant differences in the variability of beta-diversity (dispersion) and average beta-diversity (location) of the paralarval community between depth bins. Significant p-values are bolded.

| | F-statistic | p-value |
|---|---|---|
| *Variability (dispersions)* | 0.5607 | 0.7221 |
| *Average beta-diversity (location)* | 3.8352 | **0.0001** |

Table 22. Tests of significant differences in the average beta-diversity (location) of the paralarval community between pairs of depth bins. Significant p-values are bolded. Pairwise p-values are adjusted using Holm's method.

| | *0-20* | *20-40* | *40-60* | *60-80* | *80-100* |
|---|---|---|---|---|---|
| *20-40* | 0.9869 | | | | |
| *40-60* | 0.6080 | 0.5190 | | | |
| *60-80* | **0.0396** | **0.0230** | 0.1602 | | |
| *80-100* | **0.0024** | **0.0024** | 0.0624 | 0.4950 | |
| *100-130* | **0.0015** | **0.0015** | **0.0015** | 0.1036 | 0.4950 |

The CAP ordination does not show much separation between the groups, especially the 0-20 m and 20-40 m depth bins, suggesting the communities do not differ greatly (Figure 21). The 100-130 m depth bin is the most distinct group, but it still overlaps with the others. Four taxa had significant indicator power values for the depth bin grouping factor, but all had indicator power values less than 0.2, suggesting that they occurred in multiple depth bins and/or not in a majority of samples from a given depth bin (Table A3.5). Cranchiidae characterized the 80-100 m depth bin while Chiroteuthidae, Joubiniteuthidae, and Pyroteuthidae characterized 100-130 m depth bin.

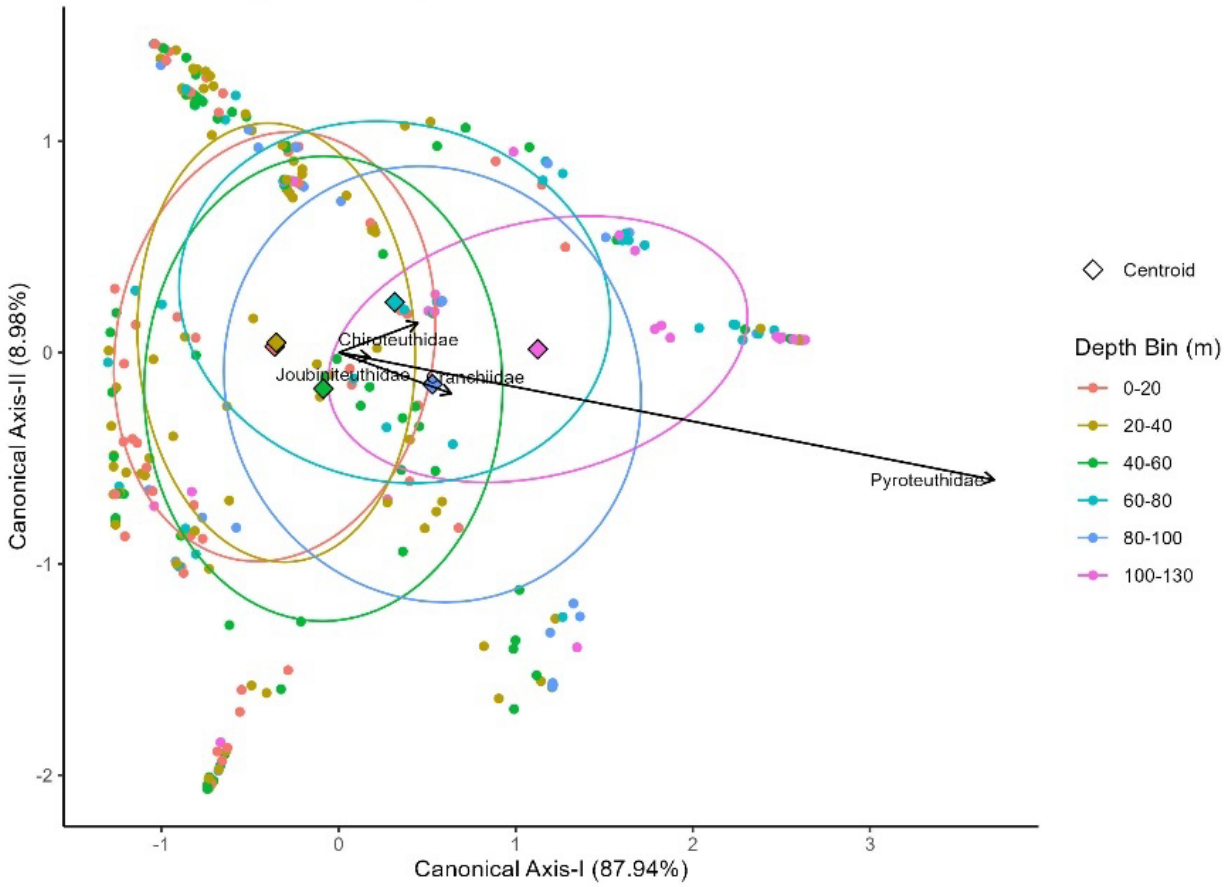


Figure 21. Canonical analysis of principal coordinates ordination for depth bin.

## 4. Discussion

This study details the abundance and distribution of twenty-one cephalopod taxa sampled from the northern GOM. Abundance was compared between time periods, regions, and depths. Additionally, modeling the abundance of three common taxa identified factors that influenced their abundance. Community composition and abundance was also analyzed to identify differences in the paralarval community among time periods, regions, and depths.

While most paralarvae could be morphologically identified to the family level, 30% could not be identified more precisely than ‘unidentified squid’ or ‘unidentified cephalopod’. This is a common issue when working with paralarvae using morphological keys. To cope with this challenge, DNA barcoding is becoming a more popular tool for identifying paralarvae (e.g., Olmos-Pérez et al., 2018; Taite et al., 2020). However, barcoding may have other issues, as it

depends on accurate identification of individuals in a database such as GenBank or BOLD. For species that have not had sequences deposited in such databases, DNA barcoding is often no more useful than morphological identification. In their study comparing morphological and molecular identification of paralarval and juvenile cephalopods, Taite et al. (2020) found that experts could accurately identify most young cephalopods, although some congeners were misidentified. A larger problem was that accurately identified sequences were not available in GenBank for all taxa (Taite et al. 2020). Other specimens morphologically identified as a certain species were found to be genetically distinct from individuals in GenBank classified as the same species (Taite et al. 2020). The findings of Taite and colleagues suggest that while DNA barcoding can be an extremely useful tool for identifying young cephalopods, especially in the absence of an expert, work needs to be done to clarify, correct (when necessary), and complete such genetic databases.

### 4.1. Abundance and Distribution

In this study, the most abundant and frequently collected taxa were Octopodidae, Enoploteuthidae, Ommastrephidae, Pyroteuthidae, and Loliginidae. These results agree with previous studies of juvenile and adult cephalopods that found these taxa to be among the most numerous families collected in the GOM (Judkins 2009, Judkins et al. 2017). Previous studies focusing on paralarvae throughout the GOM similarly found high abundance of Enoploteuthidae, Loliginidae, Ommastrephidae, and Octopodidae (Santana-Cisneros et al. 2021b, Sluis et al. 2021, García-Cordova et al. 2023).

Taxa with lower abundance and frequency of occurrence may simply be less numerous, but it is also likely that some of these taxa mainly inhabit deeper waters as paralarvae than those sampled in this study. For example, several pelagic taxa, including Bolitaenidae, Histioteuthidae, and Chiroteuthidae, were collected infrequently or not at all in this study, suggesting that their paralarvae are likely found below the sampled surface layer (0-130 m)

(Judkins et al., 2017). As juveniles and adults, these taxa are typically collected below the surface layer sampled here (e.g., below 600 m for Bolitaenidae and below 200 m for Histioteuthidae) (Judkins and Vecchione 2020). If paralarvae show the same general distribution, the low counts observed here could be expected.

Abundance of all paralarvae was highest in the early fall sampling season, but paralarvae from many taxa were present year-round, indicating that they may have extended and/or multiple spawning seasons throughout the year. Rocha et al. (2001) identified two extended spawning strategies, multiple spawning and intermittent terminal spawning, both exhibited by taxa known to occur in the GOM (e.g., Ommastrephidae and Loliginidae, respectively). Species in these families are also known to have multiple spawning seasons in a single year, e.g. *Doryteuthis pealeii* (Loliginidae) has two main spawning cohorts in the Northwest Atlantic (Hatfield and Cadrin 2002) and *Ommastrephes bartramii* (Ommastrephidae) has two main spawning cohorts in the North Pacific (Vijai et al. 2014). Although few studies have been conducted on reproductive timing of cephalopods specifically in the GOM, spawning strategies are likely similar to those found in other regions.

Of the most abundant taxa, Enoploteuthidae, Ommastrephidae, Pyroteuthidae, and Octopodidae were collected in all regions of the GOM. Loliginidae was primarily collected in the shelf regions, with only one individual collected in the slope region and none collected in the basin region. This is expected as Loliginidae are neritic as adults and spawning occurs inshore, so few paralarvae are expected to be transported offshore. Enoploteuthidae were collected more often in the shelf regions than the offshore regions, although they are considered pelagic squids. This could suggest that they spawn closer to shore or that the paralarvae are transported closer to shore by surface/shallow currents (González et al. 2005, Johnson et al. 2009, Santana-Cisneros et al. 2021a). Ommastrephidae was found at approximately the same abundance and frequency of occurrence throughout all regions except the north-central shelf region, which could also suggest inshore transport. Pyroteuthidae were found at equal

abundance and frequency of occurrence in the shelf and basin regions, but twice as often in the slope region. In other regions (e.g., northwest Atlantic), adult Pyroteuthidae were collected at offshore stations (Vecchione and Pohle 2002). If adult pyroteuthids in the GOM follow the same distribution patterns, the spatial distribution of the paralarvae may indicate dispersal due to currents. Paralarvae from the family Octopodidae were collected at much higher rates in the shelf regions than the offshore regions. As the individuals identified to this family are mainly from the genera *Octopus* and *Macrotritopus*, which are known to be neritic, benthic taxa, these findings are expected. Octopus paralarvae collected in slope and basin waters may belong to other genera, or could indicate offshore transport, which has been suggested by previous research as well as demonstrated in dispersal modeling studies (Roura et al. 2016, Santana-Cisneros et al. 2021a).

Spatial distribution in the Gulf of Mexico could also be influenced by eddies. Although the samples used in this study were not collected in eddies, it's known they can retain and transport zooplankton, including paralarvae. Eddies may have a strong effect on the transport of paralarvae with a longer paralarval duration in particular, since eddy propagation speed ranges from ~0.2 – 3.5 km/day, depending partly on latitude (Zhang et al. 2024).

### 4.2. Vertical Distribution

Most taxa had the highest abundances in the shallower depth bins. Exceptions include Chiroteuthidae, Histioteuthidae, Mastigoteuthidae, and Sepiolidae. This suggests that these taxa are located at deeper depths even at the paralarval stage. As adults they are mesopelagic taxa so it is reasonable to expect that the paralarvae do not distribute up to the surface (Quetglas et al. 2010, Judkins and Vecchione 2020, Escánez et al. 2022). Other taxa, including Argonautidae, Loliginidae, and Thysanoteuthidae were more abundant in the shallower depth bins. These are known to be epipelagic or neritic taxa, so a shallower distribution is expected for the paralarvae (Bower and Miyahara 2005, Jereb et al. 2010, Finn 2016).

Although most taxa do not show clear evidence of diel vertical migration, Pyroteuthidae, Octopodidae, Enoploteuthidae, and Onychoteuthidae may migrate vertically as paralarvae based on differing depths of maximum abundance between day and night samples. Enoploteuthid paralarvae have been shown to vertically migrate in other areas, including the Caribbean Sea and Northeast Atlantic (Diekmann et al. 2006, Castillo-Estrada et al. 2020). Juvenile and adult Enoploteuthids are also known to perform diel vertical migration (Guerra-Marrero et al. 2020, Judkins and Vecchione 2020).

This study found that enoploteuthid paralarvae were most abundant between 20-40 meters at night and 20-40 and 80-100 meters during the day (average abundance was almost equal between these two depth bins). However, as there was no significant difference in abundance between depth bins at either time of day, the abundance in all depth bins could be statistically equal. Studies conducted in the Caribbean Sea found the highest abundance of *Abralia spp.* paralarvae, a genus in the family Enoploteuthidae, in the 38-48 m depth bin (combined day and night sampling) and that of *Abralia redifieldi* in the 0-25 m depth bin at night and the 0-25 and 50-75 m depth bins during the day (Castillo-Estrada et al. 2020, García-Guillén et al. 2023). In the Gulf of California and adjacent Pacific Ocean, Enoploteuthidae paralarvae were most abundant in the 0-17 m and 17-34 m depth bins (combined day and night sampling) (Ruvalcaba-Aroche et al. 2018). In the subtropical eastern North Atlantic, enoploteuthid paralarvae were most abundant in the 10-50 m and 50-100 m depth bins at night (very similar abundances) and in the 100-150 m depth bin during the day (Diekmann et al. 2006). Ruvalcaba-Aroche et al. (2018) identified features of water masses associated with the highest abundance, including the mixed layer and thermocline, as well as warm and well-oxygenated water. Preference for features such as these could explain differences in the depth of maximum abundance in different studies.

This study found highest onychoteuthid abundance in the 0-20 m depth bin at night and the 20-40 m depth bin during the day. In the Caribbean Sea, *Onychoteuthis banksii*, a species in

the family Onychoteuthidae, was most abundant in the 0-25 m depth bin at night and the 0-25 m and 25-50 m depth bins during the day (Castillo-Estrada et al. 2020). In the subtropical eastern North Atlantic, onychoteuthid abundance was highest in the 10-50 m depth bin both at night and during the day (Diekmann et al. 2006). These results suggest that vertical distribution is similar in different regions, but there may be temporal variations. Additional fine-scale sampling in various regions would allow for more robust conclusions.

In this study, Pyroteuthidae paralarvae were most abundant in the 0-20 m depth bin at night and in the 60-80 m depth bin during the day. In the Gulf of California and adjacent Pacific Ocean, pyroteuthid paralarvae were most abundant in the 17-34 m and 34-51 m depth bins (combined day and night sampling) (Ruvalcaba-Aroche et al. 2018). In the subtropical eastern North Atlantic, they were most abundant in the 50-100 m depth bin at night and in the 100-150 m depth bin during the day (Diekmann et al. 2006). Comparing these studies suggests there may be regional variation in the depths inhabited by this group.

This study identified the highest ommastrephid abundance in the 100-130 m depth bin both at night and during the day. In the Caribbean Sea, *Ornithoteuthis antillarum,* a species in the Ommastrephidae family, was most abundant in the 0-25 m and 25-50 m depth bins at night and the 0-25 m and 50-75 m depth bins during the day (Castillo-Estrada et al. 2020). In the Gulf of California and adjacent Pacific Ocean, paralarvae of two ommastrephid species were equally abundant in three depth bins between 0 and 65 m depth (Ruvalcaba-Aroche et al. 2018). These differences may be due in part to species-specific vertical distributions. At least four species of ommastrephids are found in the GOM, including *Ornithoteuthis antillarum*, while two other species, *Dosidicus gigas* and *Sthenoteuthis oualaniensis,* are the most frequently collected ommastrephids in the Gulf of California.

It has been suggested that differing day and night abundances are due to net avoidance (Vecchione 1981), but this seems unlikely in this study as the paralarvae are far smaller than the mouth of the net and likely would not be able to swim out of the way. Additionally, to reduce net

avoidance, the nets used in these surveys were dyed black and tow speed averaged two knots. Furthermore, other studies have suggested no net avoidance in paralarvae on the basis of day and night catches lacking significant differences in abundance (Goldman and McGowan 1991, Röpke et al. 1993).

Although vertically-stratified sampling does allow for the analysis of diel vertical migration, the precision of the results are constrained by the size of the depth bins. This study used 20- or 30-meter depth bins; other studies on paralarvae have used 10 meter, 17 meter, 20 meter, 25 meter, or even larger depth bins (Goldman and McGowan 1991, Diekmann et al. 2006, Ruvalcaba-Aroche et al. 2018, Castillo-Estrada et al. 2020, García-Guillén et al. 2023). Such wide depth bins compared to the size of the animals suggests that shorter migrations cannot be precisely identified. A subset of the samples used in this study were collected over narrower depth ranges, so analyzing these samples separately could provide a more precise picture of paralarval vertical migration. Another factor influencing the results is that sites in this study were not sampled multiple times in a consecutive day-night cycle, which would allow better detail on site-specific diel vertical migration.

Other groups of diel-vertically migrating zooplankton, such as krill, migrate between 150 and 300 m from depth to the surface, although it's unlikely all individuals make the farthest migration (Riquelme-Bugueño et al. 2020). Some pteropods, a group of holoplanktonic mollusks, also vertically migrate with maximum daytime abundance between 200 and 600 m and maximum nighttime abundance between 0 and 200 m (Shedler et al. 2022). Vertical migration depends on individual swimming capability. Laboratory tests have recorded squid paralarvae swimming speeds of 16.2-48.4 $mm^{-1}$ (squids 1.8 mm DML) and 4-8 $mm^{-1}$ (squids 2.65 mm ML) (Bartol et al. 2008, Vidal et al. 2018). These values suggest that it would take cephalopod paralarvae 1.5 hours or less to cover 20 meters, assuming their vertical swimming speed is the same as their horizontal swimming speed. Such speeds suggest that paralarvae could easily migrate through the portion of the water column sampled in this study.

## 4.3. Abundance Modeling

### *4.3.1. Environmental and Temporal Variables*

This study found that water density was highly correlated with temperature, salinity, year, season, and depth bin. This is expected as density is determined mainly by salinity and temperature and temperature varies both with depth and seasonally in the GOM. The high correlation of density with year may be partly due to only having one season of 2011 samples, which led to a smaller range of temperatures being observed in that year. Due to these collinearities, an effect of salinity, season, year, or depth bin in a model may actually be due to temperature or density. For example, paralarvae may be more abundant at certain temperatures (Yoo et al. 2014, Uriarte et al. 2018), but as temperature was not included in the models, the model results may show that it is certain seasons or depth bins that have high abundance, when in fact the abundance is dependent on the temperature, and the high abundance in certain seasons or depth bins is just a side effect of the temperature dependence.

### *4.3.2. Enoploteuthidae*

Some samples had extremely high counts of Enoploteuthidae paralarvae relative to the median count, indicating that Enoploteuthidae may be patchily distributed as paralarvae. Enoploteuthid species are thought to spawn individual eggs in the plankton every few days (Young and Harman 1985, Young and Mangold 1994). This spawning strategy suggests that paralarvae may be aggregated due to physical conditions rather than the paralarvae coming from a single egg mass or spawning event. There were no obvious correlations between the count of Enoploteuthidae and any of the predictor variables. This could be due to the aggregation of samples from different cruises masking any relationships at a single time of year. It could also suggest that enoploteuthid paralarval abundance is dependent primarily on spawning times and locations rather than environmental conditions or overall time of year/location. Additionally, enoploteuthid paralarvae have a paralarval duration of approximately

1.25 months (average of 41.5 days for *Abralia trigonura*) (Bigelow 1992). This may be short enough that the effect of environmental conditions and/or time and location on paralarval abundance may not be observable.

The final model for Enoploteuthidae abundance included the variables volume filtered, year, season, diel period, and depth bin. The effect of the smooth term of volume filtered is small, but indicates that the count of Enoploteuthidae increases with more water filtered, although there is a slight dip between 250 and 400 $m^3$. Counts did not vary significantly between years. This means that enoploteuthid abundance is unlikely to be affected by year alone. Count of Enoploteuthidae is lower in the spring than the other two seasons, and higher at night than in the day. Count of Enoploteuthidae was not different between the four shallowest depth bins, while it was lower in the two deepest depth bins.

Based on the adjusted $R^2$ value, the model explains only about 6% of the variability in the count of Enoploteuthidae. This is a low value, although not unusual for biological data, suggesting that important variables may have been omitted from the model or that the structure of the model could be improved. For example, other recent studies of enoploteuthid paralarvae abundance identified other important variables than the ones identified by the model in this study. The model of enoploteuthid paralarval abundance in the northern GOM created by Sluis et al. identified distance to the Loop Current, sea surface height, and year as important variables (2021). Their results suggest that paralarval abundance is higher along fronts or eddy boundaries, potentially due to elevated primary and secondary production there (Sluis et al. 2021). In the southern GOM, García-Cordova et al. (2023) found that the variables chlorophyll-*a* concentration, depth, dissolved oxygen, sea surface temperature, water density, and geographic position were important in modeling the abundance of enoploteuthid paralarvae. Their results found highest enoploteuthid paralarval abundance in Caribbean Sea waters, but also in the frontal zones of the Yucatan Sea Water and Yucatan Upwelling Water (García-Cordova et al. 2023). Some of the variables identified as important by these studies were not included in this

study; however, their importance in other studies suggests they could be useful to consider here as well.

### *4.3.3. Octopodidae*

The octopus species included in Octopodidae for this study are merobenthic, with paralarval durations of 21 to 60 days (Vidal and Shea 2023). The eggs are attached to structures on the benthos in a den and cared for by the female (Rocha et al. 2001). As these eggs are likely to hatch around the same time, it's possible that all the paralarvae in a single sample came from a single brood, especially if the paralarvae are small (more recently hatched, meaning less time to be dispersed). There were no obvious relationships between the predictor variables and the count of Octopodidae. This could indicate that aggregating samples from different seasons, years, and spatial locations masks any relationships that may exist in a portion of the time/location. It could also indicate that the variables analyzed in this study are not the ones that paralarvae count is related to and there may be other variables that were not included which are related to paralarvae count.

The selected model for octopod abundance included the variables region, season, fluorescence, and volume filtered. The effect of volume filtered is small but positive, indicating higher counts with increased amount of water filtered, which is logical. The effect of fluorescence is also positive, suggesting that octopod abundance increases with increasing primary productivity. Paralarval octopus feed on a variety of zooplankton, which likely have higher abundance in areas of higher primary productivity. Octopod count was lower in the NC GOM shelf region and the two offshore regions than in the NW GOM shelf and WFS regions. Octopod count was higher in the early fall than the other two seasons. Based on the adjusted $R^2$, this model explains only about 12.2% of the variability in count of Octopodidae, which is fairly low, but not unreasonable for this type of data. The low adjusted $R^2$ suggests that some

important variables may have been left out of the model or that the structure of the model does not fit the population well.

Similar models produced for populations in other regions suggest other variables that may drive octopod abundance. For example, a model of paralarval octopod presence in the northeastern Atlantic (western Iberian coast) identified month, latitude, depth, upwelling index, and sea surface temperature as important predictors while a model of abundance in the same region identified year, longitude, depth, and upwelling index as important predictors (Moreno et al. 2009). Moreno et al. (2009) note that paralarval abundance and distribution are likely driven by spawning location and cross-shelf transport. In the same area, Roura et al. (2016) identified month, depth, and upwelling strength as important variables in determining octopod paralarval abundance. They observed that the distribution and abundance of paralarval octopus are likely driven by coastal upwelling and the vertical position of the paralarvae within the water column may enable transport across the shelf (Roura et al. 2016).

#### *4.3.4. Loliginidae*

One sample had an extremely high count of Loliginidae paralarvae relative to the median count, which indicates that Loliginidae may be patchily distributed as paralarvae, especially directly after hatching. Most loliginid species attach eggs to the benthos at communal spawning grounds where multiple females lay eggs (Griswold and Prezioso 1981, Shashar and Hanlon 2013). This reproductive strategy results in many paralarvae in the same area, leading to patchiness and occasional samples with high numbers of paralarvae. There were no obvious relationships between count of Loliginidae and any of the predictor variables. As discussed with the other two taxa, this may be due to sample aggregation or missing variables. It could also suggest that loliginid paralarval abundance is dependent primarily on spawning times and locations rather than environmental conditions or overall time of year/location. *Doryteuthis pealeii*, which is classified as an intermittent terminal spawner, is known to have multiple

spawning events in a year along the Atlantic coast of North America (Rocha et al. 2001, Hatfield and Cadrin 2002). Loliginid species show a range of paralarval durations from of approximately 0.25 to 2 months (e.g., 50-60 days for *Doryteuthis pealeii*) (Vidal and Shea 2023). This may be short enough that the effect of environmental conditions and/or time and location on paralarval abundance may not be visible.

The selected model for loliginid abundance is overdispersed and violates some of the model assumptions, including homogeneity of variance and normality. As such, its results should be interpreted with caution. The model included the variables year, fluorescence, transmissivity, and volume filtered. The effect of volume filtered is small but negative, indicating decreasing counts with increased amount of water filtered, which is counterintuitive. This result may be affected by the patchiness of the paralarvae, which could have an effect on the estimate of the parameter. The smoothers for fluorescence and transmissivity show small effects; fluorescence a slight increase in mid-values and transmissivity a slight increase at the low and high 90%'s, with a decrease in the mid-90s. These results suggest that loliginid abundance increases with increasing primary productivity to a certain level and that the amount of suspended material in the water has a slight effect on loliginid abundance. Count of Loliginidae was higher in samples collected in 2011 and 2012 versus 2009. This could suggest that loliginid abundance increased over the sampled years. Based on the adjusted $R^2$, this model explains about 75% of the variability in count of Loliginidae, which is quite high and may suggest model overfitting.

In the southern GOM, a study on the loliginid squid species *Doryteuthis plei* found that chlorophyll-*a*, depth, water density, and geographic position were important in explaining the variation in distribution and abundance (García-Cordova et al. 2023). They found that abundance corresponded to frontal zones of Yucatan Upwelling Water and Yucatan Sea Water, which arose from intrusions of coastal upwelling water (García-Cordova et al. 2023). In the northeastern Atlantic off western Iberia, modeling found that loliginid squid presence was driven

by month, latitude, depth, upwelling index, and sea surface temperature, while abundance was driven by upwelling index and sea surface temperature (Moreno et al. 2009). They suggest that the paralarvae are advected from the spawning location to areas with cooler, more productive upwelled waters, enhancing their survival (Moreno et al. 2009). An analysis of the distribution of loliginid species in the South Brazil Bight found that highest paralarval abundance was associated with seasonal upwelling and the intrusion of nutrient-rich waters in the nearshore, northernmost portion of the region (Araújo and Gasalla 2018).

#### *4.3.5. Model Caveats*

Based on model validation, all the models showed evidence that the model assumptions were violated to some extent, indicating that the results should be interpreted with caution. In the Enoploteuthidae and Loliginidae models, samples with outlying values were retained in the models as these are true observations of abundance. However, they are at least weakly influential on the results of the model (based on Cook's distance), and it is likely that removing them from the model would produce slightly different results that would be better fitted to the average abundance of squid. The skewness of the residuals and heterogeneity of variance suggest that a different model distribution may perform better than those tested here.

All three taxa modeled showed high percentages of samples with no paralarvae. Paralarvae are known to occur infrequently in the water column compared to other planktonic groups as well as being patchily distributed, so the high percentages of samples with no paralarvae are expected (Boletzky 2003). This indicates likely zero-inflation of the data and suggests that models with a zero-inflated distribution family, such as zero-inflated Poisson or zero-inflated negative binomial, may perform better than the standard versions of the same families, as were used here.

One feature of the sample collection that was not addressed in the models was that several samples were collected at the same site. Moran's I calculated on samples within sites

was significant and positive, indicating that the count of the taxon in samples taken at the same site was more likely to be similar than in samples collected at different sites. Addressing this correlation could be done with a generalized additive mixed model, which adds a random intercept to deal with the dependency structure (Zuur et al. 2014). Additionally, spatial and/or temporal autocorrelation structures may be required to meet the assumption of independence as the count of the taxon may be correlated between samples collected close together in space or time. For example, validation of the Octopodidae abundance model found weak but significant spatial correlation between sites, indicating that a spatial autocorrelation structure may be required to meet the independence assumption of the model.

### 4.4. Community Structure

The paralarval community in this study was grouped into four clusters of taxa. The Loliginidae-Octopodidae cluster is a logical pairing as both are neritic (Jereb et al. 2010, Rosa et al. 2019). Additionally, a study of paralarval diversity in the Southeastern Brazilian Bight similarly found that Loliginidae and Octopodidae grouped together and characterized the inner shelf (Araújo and Gasalla 2019). This study found that Pyroteuthidae and Ommastrephidae are also likely to be found together, but not with Loliginidae or Octopodidae, which is expected as they are oceanic squids. However, in the Sargasso Sea and subtropical eastern North Atlantic, these taxa were found to have opposite distributions (Diekmann and Piatkowski 2002, Diekmann et al. 2006). Interestingly, this study found that Enoploteuthidae is likely to be found alongside any of the other four taxa. As an oceanic squid, the expectation was that it would be grouped with Pyroteuthidae and Ommastrephidae. This was the case in the Southeastern Brazilian Bight, where Enoploteuthidae was frequently found alongside Ommastrephidae (Araújo and Gasalla 2019). In the Sargasso Sea, Enoploteuthidae was more commonly found alongside Pyroteuthidae (Diekmann and Piatkowski 2002).

In this study, Onychoteuthidae and Cranchiidae were both more likely to be found alongside Ommastrephidae and Pyroteuthidae, while in the Sargasso Sea, both taxa were likely to be found alongside Pyroteuthidae, but not Ommastrephidae (Diekmann and Piatkowski 2002). In the subtropical eastern North Atlantic, Onychoteuthidae was likely to be found alongside Ommastrephidae but not Pyroteuthidae, while Cranchiidae was found with both, based on subfamily (Diekmann et al. 2006). These differences suggest that assemblages may be driven by different factors in different regions or that differences in the individual species found in a given region may affect assemblages at the family level.

This study found that the paralarval community differed between different years, being more similar between 2009 and 2012 than either of those years and 2011, which could indicate the differences in sampling effort among years. This result may also reflect seasonal changes in the paralarval community, as 2011 was only sampled in one season. However, it could also represent a temporary change in the paralarval community due to the *Deepwater Horizon* oil spill. Variations in the overall zooplankton assemblage were identified in May and June 2010, but did not persist significantly in July 2010 (Carassou et al. 2014). If the cephalopod community showed a similar trajectory, the variability identified in the 2011 community is unlikely to be due to the oil spill. Additional sampling of the zooplankton community also showed no significant differences in beta-diversity between summer 2010 and other years (Daly et al. 2021). However, cephalopods often show large annual fluctuations in population abundance, due in part to sensitivity to environmental variation, so the effects of the *Deepwater Horizon* oil spill could persist in the cephalopod community longer than in the zooplankton community as a whole (Rodhouse et al. 2014).

The paralarval community also varied between seasons. This variability is likely due to different spawning seasons and paralarval durations among the different taxa. Despite these differences, the analysis also suggested that beta-diversity is more similar between winter and early fall than winter and spring or early fall and spring, while beta-diversity in the spring was

more similar to winter. This could reflect lower sampling effort in the spring or it could suggest a gap in spawning that results in differences in the taxa present and their abundances between spring and early fall. Seasonal differences in assemblages were also identified in the Gulf of California, which was hypothesized to be due to changes in temperature and productivity (De Silva-Dávila et al. 2015).

The beta-diversity of the paralarval community also varied between daytime and nighttime samples. This could be due to different taxa migrating into or out of the sampled portion of the water column as the day progresses. Diel differences in the zooplankton community, both in abundance and taxa composition, have been recorded in a range of locations (Nakajima et al. 2008, Kimmel et al. 2010). As members of the plankton community, cephalopod paralarvae likely follow similar patterns.

The composition and abundance of the paralarval community in the GOM slope and GOM basin regions was not different. The beta-diversity of the three shelf regions was different from the slope and basin regions, which is expected as Loliginidae and Octopodidae are influential taxa that are mainly found on the shelf (Jereb et al. 2010, Rosa et al. 2019). The beta-diversity also varied between the NC GOM shelf and the WFS, but not between the NW GOM shelf and the NC GOM shelf or the NW GOM shelf and the WFS. This indicates that although the NC GOM shelf and the WFS are contiguous, the paralarval communities in the regions are different in either taxa composition and/or abundance. The beta-diversity does not differ between the other pairs of shelf regions. However, this study does not identify the causes of these differences or lack thereof. Modeling studies have identified cross-shelf dispersal in ichthyoplankton (Lara-Hernández et al. 2019, Vasbinder et al. 2023) and cephalopod paralarvae may experience similar dispersal patterns. These studies identify areas of the shelf that are more connected than others, which could result in a more similar plankton community, including paralarvae, in those areas.

The difference in the paralarval community between depths suggests that the taxa found at the surface are either not found in the deeper bins or found in much lower abundance, and vice versa. The vertical distribution analysis in this study explains which taxon/taxa may be influencing this difference, most likely Loliginidae. The second depth bin (20-40 m) also had a different beta-diversity from the three deepest depth bins which suggests it has a similar assemblage to the shallowest depth bin. The shift in the paralarval community likely occurs over the 40-60 m depth bin, as it does not have different beta-diversity from any of the depth bins except the deepest one. Further investigation could suggest causes for this, such as physical oceanographic features including the thermocline. The three deepest depth bins do not vary in taxa composition and abundance, which indicates that below 60 m (at least down to 130 m) the paralarval community is similar.

**5. Conclusions**

This study examined different aspects of cephalopod paralarval ecology in the northern GOM. Paralarvae were collected in fewer than two-thirds of the samples, with much lower frequency of occurrence for individual taxa. Enoploteuthidae, Loliginidae, Octopodidae, Ommastrephidae, and Pyroteuthidae were identified as the most abundant taxa. Most taxa did not show significant differences in abundance between years, seasons, or regions. The majority of abundant taxa were found throughout the upper water column (0-130 m). Some taxa, including Pyroteuthidae, showed evidence of diel vertical migration. Most were found deeper during the day than at night, on average. The abundance models created in this study all had some degree of model misspecification and violation of model assumptions. However, they do suggest that different factors influence the abundance of the different taxa, although there are some similarities. Year was important for the squid taxa, while fluorescence was important for the neritic taxa. Finally, the analysis of community structure identified differences in the composition and abundance of the paralarval community between years, seasons, diel periods,

regions, and depth bins, although some similar communities were identified, including the GOM slope and GOM basin, the shallowest depth bins, and the deepest depth bins. However, this study, like many that focus on paralarvae, is strongly impacted by extreme variability in abundance (patchiness), which affected the results, including average abundances and the existence of significant differences in abundance.

The results of this study provide useful information about the cephalopod paralarval community in the northern GOM. They can be compared with paralarvae abundance and distribution information in other regions to understand differences in the communities and identify any similarities. These results can also be reviewed in conjunction with studies on juvenile and adult cephalopods in the GOM to make connections between different life cycle stages. The models also suggest directions for future modeling efforts in terms of model type and the variables that may be important to be included. These data also provide a baseline of cephalopod paralarvae abundance and distribution to compare to previously collected samples or future samples. The scale of the data in space and time also provides suggestions on where or when to target certain taxa.

## References


Allard, J., J. V Clarke III, and B. D. Keim. 2016. Spatial and Temporal Patterns of In Situ Sea Surface Temperatures within the Gulf of Mexico from 1901-2010. American Journal of Climate Change 5:314–343.

Allee, R., R. Finkbeiner, R. Gould, D.-S. Ko, D. Lary, J. Kurtz, C. Madden, and K. Goodin. 2012. Coastal and Marine Ecological Classification Standard (CMECS) Application for Sea Surface Salinity in the Northern Gulf of Mexico. https://gulfatlas.noaa.gov/.

Anderson, R. C., J. B. Wood, and J. A. Mather. 2008. *Octopus vulgaris* in the Caribbean is a specializing generalist. Marine Ecology Progress Series 371:199–202.

Araújo, C. C. de, and M. A. Gasalla. 2018. Distribution patterns of loliginid squid paralarvae in relation to the oceanographic features off the South Brazil Bight (22°–25°S). Fisheries Oceanography 27:63–75.

Araújo, C. C. de, and M. A. Gasalla. 2019. Biodiversity of cephalopod early-life stages across the Southeastern Brazilian Bight: spatio-temporal patterns in taxonomic richness. Marine Biodiversity 49:2429–2443.

Arkhipkin, A. I. 2004. Diversity in growth and longevity in short-lived animals: Squid of the suborder Oegopsina. Marine and Freshwater Research 55:341–355.

Bartol, I. K., P. S. Krueger, J. T. Thompson, and W. J. Stewart. 2008. Swimming dynamics and propulsive efficiency of squids throughout ontogeny. Integrative and Comparative Biology 48:720–733.

Bigelow, K. A. 1992. Age and growth in paralarvae of the mesopelagic squid *Abralia trigonura* based on daily growth increments in statoliths. Marine Ecology Progress Series 82:31–40.

Boletzky, S. v. 2003. Biology of Early Life Stages in Cephalopod Molluscs. Pages 143–203 *in* A. J. Southward, P. A. Tyler, C. M. Young, and L. A. Fuiman, editors. Advances in Marine Biology. Elsevier Science.

Bower, J. R., and K. Miyahara. 2005. The diamond squid (*Thysanoteuthis rhombus*): A review of the fishery and recent research in Japan. Fisheries Research 73:1–11.

Bower, J. R., and S. Takagi. 2004. Summer vertical distribution of paralarval gonatid squids in the northeast Pacific. Journal of Plankton Research 26:851–857.

Candela, J., J. Sheinbaum, J. Ochoa, A. Badan, and R. Leben. 2002. The potential vorticity flux through the Yucatan Channel and the Loop Current in the Gulf of Mexico. Geophysical Research Letters 29:16--1-16–4.

Carassou, L., F. J. Hernandez, and W. M. Graham. 2014. Change and recovery of coastal mesozooplankton community structure during the Deepwater Horizon oil spill. Environmental Research Letters 9:124003.

Castillo-Estrada, G., R. De Silva-Dávila, L. Carrillo, L. Vásquez-Yeomans, C. A. Silva-Segundo, L. Avilés-Diáz, and U. Markaida. 2020. Horizontal and vertical distribution of cephalopod paralarvae in the Mesoamerican Barrier Reef System. Journal of the Marine Biological Association of the United Kingdom 100:927–937.

Coll, M., J. Navarro, R. J. Olson, and V. Christensen. 2013. Assessing the trophic position and ecological role of squids in marine ecosystems by means of food-web models. Deep-Sea Research Part II: Topical Studies in Oceanography 95:21–36.

Condie, S., and R. Condie. 2016. Retention of plankton within ocean eddies. Global Ecology and Biogeography 25:1264–1277.

Daly, K. L., A. Remsen, D. M. Outram, H. Broadbent, K. Kramer, and K. Dubickas. 2021. Resilience of the zooplankton community in the northeast Gulf of Mexico during and after the Deepwater Horizon oil spill. Marine Pollution Bulletin 163:111882.

Damien, P., J. Sheinbaum, O. Pasqueron De Fommervault, J. Jouanno, L. Linacre, and O. Duteil. 2021. Do Loop Current eddies stimulate productivity in the Gulf of Mexico? Biogeosciences 18:4281–4303.

DeRuiter, S. 2019. Collinearity and Multicollinearity. https://stacyderuiter.github.io/s245-notes-bookdown/collinearity-and-multicollinearity.html.

Diekmann, R., W. Nellen, and U. Piatkowski. 2006. A multivariate analysis of larval fish and paralarval cephalopod assemblages at Great Meteor Seamount. Deep-Sea Research Part I: Oceanographic Research Papers 53:1635–1657.

Diekmann, R., and U. Piatkowski. 2002. Early life stages of cephalopods in the Sargasso Sea: distribution and diversity relative to hydrographic conditions. Marine Biology 141:123–130.

Elliott, B. A. 1982. Anticyclonic Rings in the Gulf of Mexico. Journal of Physical Oceanography 12:1292–1309.

Escánez, A., Á. Guerra, R. Riera, A. Ariza, Á. F. González, and N. Aguilar de Soto. 2022. New contribution to the knowledge of the mesopelagic cephalopod community off the western Canary Islands slope. Regional Studies in Marine Science 55:102572.

FAO. 2020. The State of World Fisheries and Aquaculture 2020. Sustainability in action. Rome.

FAO. 2022. The State of World Fisheries and Aquaculture 2022. Towards Blue Transformation. Rome.

Fasiolo, M., R. Nedellec, Y. Goude, and S. N. Wood. 2020. Scalable visualization methods for modern generalized additive models. Journal of Computational and Graphical Statistics 29:78–86.

Fernández-Álvarez, F., A. Machordom, R. García-Jiménez, C. A. Salinas-Zavala, and R. Villanueva. 2018. Predatory flying squids are detritivores during their early planktonic life. Scientific Reports 8:3440.

Finn, J. K. 2016. Family Argonautidae Tryon, 1879. Pages 228–237 *in* P. Jereb, C. F. E. Roper, M. D. Norman, and J. K. Finn, editors. Cephalopods of the World. An annotated and illustrated catalogue of cephalopod species known to date. Second. FAO, Rome.

Fox, J., and S. Weisberg. 2018. An R Companion to Applied Regression. Third edition. Sage Publications, Inc, Thousand Oaks, CA.

Garcia, H. 2014. Dissolved Oxygen - Climatological Mean In Gulf of Mexico Data Atlas. https://gulfatlas.noaa.gov/.

García-Cordova, E. A., R. De Silva-Dávila, I. Velázquez-Abunader, J. Q. García-Maldonado, and P.-L. Ardisson. 2023. Distribution of squid paralarvae and related oceanographic features in the eastern Campeche Bank, Mexico. Journal of Plankton Research 45:278–290.

García-Guillén, R. M., R. De Silva-Dávila, C. A. Silva-Segundo, J. F. Domínguez-Contreras, L. Carrillo, and L. Vásquez-Yeomans. 2023. Horizontal and vertical distribution of *Abralia spp.* paralarvae in the Caribbean Sea with morphological and molecular notes. Bulletin of Marine Science 99:143–168.

Goldman, D. A., and M. F. McGowan. 1991. Distribution and abundance of Ommastrephid squid paralarvae off the Florida Keys in August 1989. Bulletin of Marine Science 49:614–622.

González, A. F., J. Otero, A. Guerra, R. Prego, F. J. Rocha, and A. W. Dale. 2005. Distribution of common octopus and common squid paralarvae in a wind-driven upwelling area (Ria of Vigo, northwestern Spain). Journal of Plankton Research 27:271–277.

Griswold, C. A., and J. Prezioso. 1981. In situ observations on reproductive behavior of the long-finned squid, *Loligo pealei*. Fishery Bulletin 78:945–947.

Guarneros-Narváez, P. V., R. Rodríguez-Canul, R. De Silva-Dávila, J. A. Zamora-Briseño, M. Améndola-Pimenta, A. J. Souza, U. Ordoñez, and I. Velázquez-Abunader. 2022. Loliginid paralarvae from the Southeastern Gulf of Mexico: Abundance, distribution, and genetic structure. Frontiers in Marine Science 9:941908.

Guerra-Marrero, A., V. Hernández-García, A. Sarmiento-Lezcano, D. Jiménez-Alvarado, A. Santana-Del Pino, and J. J. Castro. 2020. Migratory patterns, vertical distributions and diets of *Abralia veranyi* and *Abraliopsis morisii* (Cephalopoda: Enoploteuthidae) in the eastern North Atlantic. Journal of Molluscan Studies 86:27–34.

Haimovici, M., U. Piatkowski, and R. A. dos Santos. 2002. Cephalopod paralarvae around tropical seamounts and oceanic islands off the north-eastern coast of Brazil. Bulletin of Marine Science 71:313–330.

Hatfield, E. M. C., and S. X. Cadrin. 2002. Geographic and temporal patterns in size and maturity of the longfin inshore squid (*Loligo pealeii*) off the northeastern United States. Fisheries Bulletin 100:200–213.

Herve, M. 2023. _RVAideMemoire: Testing and Plotting Procedures for Biostatistics_.

Hidaka, K., K. Kawaguchi, M. Murakami, and M. Takahashi. 2001. Downward transport of organic carbon by diel migratory micronekton in the western equatorial Pacific: its quantitative and qualitative importance. Deep-Sea Research I 48:1923–1939.

Hoving, H. J. T., J. A. A. Perez, K. S. R. Bolstad, H. E. Braid, A. B. Evans, D. Fuchs, H. Judkins, J. T. Kelly, J. E. A. R. Marian, R. Nakajima, U. Piatkowski, A. Reid, M. Vecchione, and J. C. C. Xavier. 2014. The study of deep-sea cephalopods. Pages 235–359 Advances in Marine Biology.

Ieno, E. N., A. F. Zuur, R. A. Fuller, T. Piersma, and A. A. Saveliev. 2014. Poisson GAMM applied on ruddy turnstone data. Pages 99–114 *in* A. F. Zuur, A. A. Saveliev, and E. N. Ieno, editors. A Beginner's Guide to Generalized Additive Mixed Models with R. Highland Statistics Ltd., Newburgh, UK.

Jereb, P., M. Vecchione, and C. F. E. Roper. 2010. Family Loliginidae Lesueur, 1821. Pages 38–117 *in* P. Jereb and C. F. E. Roper, editors. Cephalopods of the world. An annotated and illustrated catalogue of cephalopod species known to date. Second. FAO, Rome.

Johnson, D. R., H. M. Perry, J. Lyczkowski-Shultz, and D. Hanisko. 2009. Red Snapper Larval Transport in the Northern Gulf of Mexico. Transactions of the American Fisheries Society 138:458–470.

Jones, N. J. E., and C. A. Richardson. 2010. Laboratory culture, growth, and the life cycle of the little cuttlefish *Sepiola atlantica* (Cephalopoda: Sepiolidae). Journal of Shellfish Research 29:241–246.

Judkins, H. L. 2009. Cephalopods of the Broad Caribbean: Distribution, Abundance, and Ecological Importance. Doctor of Philosophy, University of South Florida.

Judkins, H., and M. Vecchione. 2020. Vertical Distribution Patterns of Cephalopods in the Northern Gulf of Mexico. Frontiers in Marine Science 7:1–18.

Judkins, H., M. Vecchione, A. Cook, and T. Sutton. 2017. Diversity of midwater cephalopods in the northern Gulf of Mexico: comparison of two collecting methods. Marine Biodiversity 47:647–657.

Kimmel, D. G., W. C. Boicourt, J. J. Pierson, M. R. Roman, and X. Zhang. 2010. The vertical distribution and diel variability of mesozooplankton biomass, abundance and size in response to hypoxia in the northern Gulf of Mexico USA. Journal of Plankton Research 32:1185–1202.

Kindt, R., and R. Coe. 2005. _Tree diversity analysis. A manual and software for common statistical methods for ecological and biodiversity studies_. World Agroforestry Centre (ICRAF).

Kourafalou, V., Y. Androulidakis, M. Le Hénaff, and H. S. Kang. 2017. The Dynamics of Cuba Anticyclones (CubANs) and Interaction With the Loop Current/Florida Current System. Journal of Geophysical Research: Oceans 122:7897–7923.

Kuhn, M. 2008. Building Predictive Models in R Using the caret Package. Journal of Statistical Software 28:1–26.

Lara-Hernández, J. A., J. Zavala-Hidalgo, L. Sanvicente-Añorve, and P. Briones-Fourzán. 2019. Connectivity and larval dispersal pathways of *Panulirus argus* in the Gulf of Mexico: A numerical study. Journal of Sea Research 155:101814.

Leben, R. R. 2005. Altimeter-Derived Loop Current Metrics. Geophysical Monograph Series:1–22.

Liu, H., and M. Dagg. 2003. Interactions between nutrients, phytoplankton growth, and micro- and mesozooplankton grazing in the plume of the Mississippi River. Marine Ecology Progress Series 258:31–42.

Lohrenz, S. E., M. J. Daggt, and T. E. Whitledge. 1990. Enhanced primary production at the plume/oceanic interface of the Mississippi River. Continental Shelf Research 10:639–664.

Luis, R. R., L. A. S. Alcalde, O. Gutiérrez-Benítez, L. Jiménez-Badillo, and C. A. Villegas Sánchez. 2020. Octopus fisheries on the Veracruz Reef system of the Gulf of Mexico: Tendencies and fluctuations. Regional Studies in Marine Science 39:101400.

Martínez-Soler, E., J. Gómez-Gutiérrez, R. De Silva-Dávila, E. González-Rodríguez, and O. Aburto-Oropeza. 2021. Cephalopod paralarval species richness, abundance and size structure during the 2014-2017 anomalous warm period in the southern Gulf of California. Journal of Plankton Research 43:224–243.

Maul, G. A. 1977. The annual cycle of the Gulf Loop Current Part I: Observations during a one-year time series. Journal of Marine Research 35:29–47.

Metz, M., C. Haug, and J. T. Haug. 2015. Autofluorescence microscopy as a method for the documentation of cephalopod paralarvae and juveniles. Ruthenica 25:105–116.

Miller, D. L., E. Pedersen, and G. L. Simpson. 2016, August 6. GAMs: Model Selection. https://eric-pedersen.github.io/mgcv-esa-workshop/slides/03-model-selection.html#/.

Moreno, A., A. dos Santos, U. Piatkowski, A. M. P. Santos, and H. Cabral. 2009. Distribution of cephalopod paralarvae in relation to the regional oceanography of the western Iberia. Journal of Plankton Research 31:73–91.

Morey, S. L., W. W. Schroeder, J. J. O'Brien, and J. Zavala-Hidalgo. 2003. The annual cycle of riverine influence in the eastern Gulf of Mexico basin. Geophysical Research Letters 30:5-1-5–4.

Muhling, B. A., R. H. Smith, L. Vásquez-Yeomans, J. T. Lamkin, E. M. Johns, L. Carrillo, E. Sosa-Cordero, and E. Malca. 2013. Larval fish assemblages and mesoscale oceanographic structure along the Mesoamerican Barrier Reef System. Fisheries Oceanography 22:409–428.

Nakajima, R., T. Yoshida, B. H. R. Othman, and T. Toda. 2008. Diel variation in abundance, biomass and size composition of zooplankton community over a coral-reef in Redang Island, Malaysia. Plankton and Benthos Research 3:216–226.

National Park Service. 2022, February 10. Mississippi River Facts - Mississippi National River and Recreation Area (U.S. National Park Service). https://www.nps.gov/miss/riverfacts.htm.

NOAA Fisheries. 2020. Species Directory. https://www.fisheries.noaa.gov/species-directory.

NOAA Fisheries. 2022, July. Landings. https://www.fisheries.noaa.gov/foss/.

Oksanen, J., G. Simpson, F. Blanchet, R. Kindt, P. Legendre, P. Minchin, R. O'Hara, P. Solymos, M. Stevens, E. Szoecs, H. Wagner, M. Barbour, M. Bedward, B. Bolker, D. Borcard, G. Carvalho, M. Chirico, M. De Caceres, S. Durand, H. Evangelista, R. FitzJohn, M. Friendly, B. Furneaux, G. Hannigan, M. Hill, L. Lahti, D. McGlinn, M. Ouellette, C. E. Ribeiro, T. Smith, A. Stier, C. Ter Braak, and J. Weedon. 2022. _vegan: Community Ecology Package_.

Olmos-Pérez, L., Á. Roura, G. J. Pierce, and Á. F. González. 2018. Sepiolid paralarval diversity in a regional upwelling area of the NE Atlantic. Hydrobiologia 808:57–70.

Ortner, P. B., L. C. Hill, and S. R. Cummings. 1989. Zooplankton community structure and copepod species composition in the northern Gulf of Mexico. Continental Shelf Research 9:387–402.

Ospina-Alvarez, A., S. De Juan, P. Pita, G. B. Ainsworth, F. L. Matos, C. Pita, and S. Villasante. 2022. A network analysis of global cephalopod trade. Nature Scientific Reports 12:322.

Paradis, E., and K. Schliep. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35:526–528.

Pasqueron De Fommervault, O., P. Perez-Brunius, P. Damien, V. F. Camacho-Ibar, and J. Sheinbaum. 2017. Temporal variability of chlorophyll distribution in the Gulf of Mexico: bio-optical data from profiling floats. Biogeosciences 14:5647–5662.

Piatkowski, U., G. J. Pierce, and M. Morais da Cunha. 2001. Impact of cephalopods in the food chain and their interaction with the environment and fisheries: an overview. Fisheries Research 52:5–10.

Portner, E. J., U. Markaida, C. J. Robinson, and W. F. Gilly. 2020. Trophic ecology of Humboldt squid, *Dosidicus gigas*, in conjunction with body size and climatic variability in the Gulf of California, Mexico. Limnology and Oceanography 65:732–748.

Posit team. 2023. RStudio: Integrated Development Environment for R. Posit Software, PBC, Boston, MA.

Power, J. H. 1989. Sink or Swim: Growth Dynamics and Zooplankton Hydromechanics. The American Naturalist 133:706–721.

Quensen, J., G. Simpson, and J. Oksanen. 2024. _ggordiplots: Make "ggplot2" Versions of Vegan's Ordiplots_.

Quetglas, A., A. de Mesa, F. Ordines, and A. Grau. 2010. Life history of the deep-sea cephalopod family Histioteuthidae in the western Mediterranean. Deep-Sea Research I 57:999–1008.

R Core Team. 2023. _R: A Language and Environment for Statistical Computing_. R Foundation for Statistical Computing, Vienna, Austria.

Ramos-Castillejos, J. E., C. A. Salinas-Zavala, S. Camarillo-Coop, and L. M. Enríquez-Paredes. 2010. Paralarvae of the jumbo squid, *Dosidicus gigas*. Invertebrate Biology 129:172–183.

Richards, T. M., T. T. Sutton, M. S. Woodstock, H. Judkins, and R. J. David Wells. 2023. Body size, depth of occurrence, and local oceanography shape trophic structure in a diverse deep-pelagic micronekton assemblage. Progress in Oceanography 213:102998.

Riquelme-Bugueño, R., I. Pérez-Santos, N. Alegría, C. A. Vargas, M. A. Urbina, and R. Escribano. 2020. Diel vertical migration into anoxic and high-pCO 2 waters: acoustic and net-based krill observations in the Humboldt Current. Nature Scientific Reports 10:17181.

Roberts, D. W. 2023. _labdsv: Ordination and Multivariate Analysis for Ecology_.

Rocha, F., Á. Guerra, and Á. F. González. 2001. A review of reproductive strategies in cephalopods. Biological Reviews 76:291–304.

Rodhouse, P. G. K. 2013. Role of squid in the Southern Ocean pelagic ecosystem and the possible consequences of climate change. Deep-Sea Research II 95:129–138.

Rodhouse, P. G. K., G. J. Pierce, O. C. Nichols, W. H. H. Sauer, A. I. Arkhipkin, V. V. Laptikhovsky, M. R. Lipiński, J. E. Ramos, M. Gras, H. Kidokoro, K. Sadayasu, J. Pereira, E. Lefkaditou, C. Pita, M. Gasalla, M. Haimovici, M. Sakai, and N. Downey. 2014. Environmental effects on cephalopod population dynamics: Implications for management of fisheries. Pages 99–233 Advances in Marine Biology. Academic Press.

Röpke, A., W. Nellen, and U. Piatkowski. 1993. A comparative study on the influence of the pycnocline on the vertical distribution of fish larvae and cephalopod paralarvae in three ecologically different areas of the Arabian Sea. Deep-Sea Research II 40:801–819.

Rosa, R., V. Pissarra, F. O. Borges, J. Xavier, I. Gleadall, A. Golikov, G. Bello, L. Morais, F. Lishchenko, Á. Roura, H. Judkins, C. M. Ibáñez, U. Piatkowski, M. Vecchione, and R. Villanueva. 2019. Global patterns of species richness in coastal cephalopods. Frontiers in Marine Science 6:469.

Rosas-Luis, R., J. Navarro, P. Sánchez, and J. L. del Río. 2016. Assessing the trophic ecology of three sympatric squid in the marine ecosystem off the Patagonian Shelf by combining stomach content and stable isotopic analyses. Marine Biology Research 12:402–411.

Roura, Á., X. A. Álvarez-Salgado, Á. F. González, M. Gregori, G. Rosón, J. Otero, and Á. Guerra. 2016. Life strategies of cephalopod paralarvae in a coastal upwelling system (NW Iberian Peninsula): insights from zooplankton community and spatio-temporal analyses. Fisheries Oceanography 25:241–258.

Ruvalcaba-Aroche, E. D., L. Sánchez-Velasco, E. Beier, V. M. Godínez, E. D. Barton, and Ma. R. Pacheco. 2018. Effects of mesoscale structures on the distribution of cephalopod paralarvae in the Gulf of California and adjacent Pacific. Deep-Sea Research Part I: Oceanographic Research Papers 131:62–74.

Santana-Cisneros, M. L., P.-L. Ardisson, Á. F. González, I. Mariño-Tapia, M. Cahuich-López, L. E. Ángeles-González, U. Ordoñez-López, and I. Velázquez-Abunader. 2021a. Dispersal modeling of octopoda paralarvae in the Gulf of Mexico. Fisheries Oceanography 30:726–739.

Santana-Cisneros, M. L., R. Rodríguez-Canul, J. A. Zamora-Briseño, M. Améndola-Pimenta, R. De Silva-Dávila, U. Ordóñez-López, I. Velázquez-Abunader, and P.-L. Ardisson. 2021b. Morphological and molecular identification of Octopoda (Mollusca: Cephalopoda) paralarvae from the southern Gulf of Mexico. Bulletin of Marine Science 97:281–304.

Schiller, R. V., and V. H. Kourafalou. 2014. Loop Current Impact on the Transport of Mississippi River Waters. Journal of Coastal Research 298:1287–1306.

Shashar, N., and R. T. Hanlon. 2013. Spawning behavior dynamics at communal egg beds in the squid *Doryteuthis (Loligo) pealeii*. Journal of Experimental Marine Biology and Ecology 447:65–74.

Shedler, S., B. Seibel, M. Vecchione, D. Griffin, and H. Judkins. 2022. Abundance and distribution of large calcareous thecosome pteropods in the northern Gulf of Mexico. American Malacological Bulletin 39:1–11.

da Silva, C. E., and R. M. Castelao. 2018. Mississippi River Plume Variability in the Gulf of Mexico From SMAP and MODIS-Aqua Observations. Journal of Geophysical Research: Oceans 123:6620–6638.

De Silva-Dávila, R., C. Franco-Gordo, F. G. Hochberg, E. Godínez-Domínguez, R. Avendaño-Ibarra, J. Gómez-Gutiérrez, and C. J. Robinson. 2015. Cephalopod paralarval assemblages in the Gulf of California during 2004−2007. Marine Ecology Progress Series 520:123–141.

Sluis, M. Z., H. Judkins, M. A. Dance, M. Vecchione, M. Cornic, T. Sutton, and J. R. Rooker. 2021. Taxonomic composition, abundance and habitat associations of squid paralarvae in the northern Gulf of Mexico. Deep Sea Research Part I: Oceanographic Research Papers 174:103572.

Southeast Area Monitoring and Assessment Program. 2016. SEAMAP Operations Manual for Trawl and Plankton Surveys. Ocean Springs.

Staudinger, M. D., V. H. Dimkovikj, C. A. M. France, E. Jorgensen, H. Judkins, A. Lindgren, E. K. Shea, and M. Vecchione. 2019. Trophic ecology of the deep-sea cephalopod assemblage near Bear Seamount in the Northwest Atlantic Ocean. Marine Ecology Progress Series 629:67–86.

Sturges, W., and R. Leben. 2000. Frequency of Ring Separations from the Loop Current in the Gulf of Mexico: A Revised Estimate. Journal of Physical Oceanography 30:1814–1819.

Sutton, T. T., R. J. Milligan, K. Daly, K. M. Boswell, A. B. Cook, M. Cornic, T. Frank, K. Frasier, D. Hahn, F. Hernandez, J. Hildebrand, C. Hu, M. W. Johnston, S. B. Joye, H. Judkins, J. A. Moore, S. A. Murawski, N. M. Pruzinsky, J. A. Quinlan, A. Remsen, K. L. Robinson, I. C. Romero, J. R. Rooker, M. Vecchione, and R. J. D. Wells. 2022. The Open-Ocean Gulf of Mexico After Deepwater Horizon: Synthesis of a Decade of Research. Frontiers in Marine Science 9:753391.

Sweeney, M. J., C. F. E. Roper, K. M. Mangold, M. R. Clarke, and S. V Boletzky. 1992. “Larval” and Juvenile Cephalopods: A Manual for Their Identification. Smithsonian Contributions to Zoology 513:1–282.

Taite, M., M. Vecchione, S. Fennell, and A. L. Allcock. 2020. Paralarval and juvenile cephalopods within warm-core eddies in the North Atlantic. Bulletin of Marine Science 96:235–261.

Uriarte, I., J. Hernández, J. Dorner, K. Paschke, A. Farias, E. Crovetto, and C. Rosas. 2010. Rearing and growth of the octopus *Robsonella fontaniana* (Cephalopoda: Octopodidae) from planktonic hatchlings to benthic juveniles. Biological Bulletin 218:200–210.

Uriarte, I., C. Rosas, V. Espinoza, J. Hernández, and A. Farías. 2018. Thermal tolerance of paralarvae of Patagonian red octopus *Enteroctopus megalocyathus*. Aquaculture Research 49:2119–2127.

Vasbinder, K., C. H. Ainsworth, Y. Liu, and R. H. Weisberg. 2023. Gulf of Mexico larval dispersal: Combining concurrent sampling, behavioral, and hydrodynamic data to inform end-to-end modeling efforts through a Lagrangian dispersal model. Deep-Sea Research II 211:105323.

Vecchione, M. 1981. Aspects of the early life history of *Loligo pealei* (Cephalopoda: Myopsida). Journal of Shellfish Research 1:171–180.

Vecchione, M. 1999. Extraordinary abundance of squid paralarvae in the tropical eastern Pacific Ocean during El Niño of 1987. Fisheries Bulletin 97:1025–1030.

Vecchione, M., and G. Pohle. 2002. Midwater cephalopods in the Western North Atlantic Ocean off Novia Scotia. Bulletin of Marine Science 71:883–892.

Vecchione, M., C. F. E. Roper, M. J. Sweeney, and C. C. Lu. 2001. Distribution, Relative Abundance and Developmental Morphology of Paralarval Cephalopods in the Western North Atlantic Ocean. Seattle.

Vidal, E. A. G., and E. K. Shea. 2023. Cephalopod ontogeny and life cycle patterns. Frontiers in Marine Science 10:1162735.

Vidal, E. A. G., L. D. Zeidberg, and E. J. Buskey. 2018. Development of swimming abilities in squid paralarvae: Behavioral and ecological implications for dispersal. Frontiers in Physiology 9:954.

Vijai, D., M. Sakai, Y. Kamei, and Y. Sakurai. 2014. Spawning pattern of the neon flying squid *Ommastrephes bartramii* (Cephalopoda: Oegopsida) around the Hawaiian Islands. Scientia Marina 78:511–519.

Villanueva, R., E. A. G. Vidal, F. Á. Fernández-Álvarez, and J. Nabhitabhata. 2016. Early Mode of Life and Hatchling Size in Cephalopod Molluscs: Influence on the Species Distributional Ranges. PLoS ONE 11:e0165334.

Voss, G., and T. Brakoniecki. 1985. Squid Resources of the Gulf of Mexico and Southeast Atlantic Coasts of the United States. NAFO Scientific Council Studies 9:27–37.

Voss, G. L. 1954. Cephalopoda of the Gulf of Mexico. Pages 1–604 *in* P. S. Galtsoff, editor. Gulf of Mexico: Its Origin, Waters, and Marine Life. United States Government Printing Office, Washington, D.C.

Watanabe, H., T. Kubodera, M. Moku, and K. Kawaguchi. 2006. Diel vertical migration of squid in the warm core ring and cold water masses in the transition region of the western North Pacific. Marine Ecology Progress Series 315:187–197.

Wells, M. J., and A. Clarke. 1996. Energetics: the costs of living and reproducing for an individual cephalopod. Philosophical Transactions of the Royal Society B: Biological Sciences 351:1083–1104.

Wilkinson, T. A. C., E. Wiken, J. Bezaury Creel, T. F. Hourigan, T. Agardy, H. Herrmann, L. Janishevski, C. Madden, L. Morgan, and M. Padilla. 2009. Marine Ecoregions of North America. Commission for Environmental Cooperation, Montreal.

Wood, S. N. 2017. Generalized Additive Models. Second edition. Chapman and Hall/CRC, New York.

Yoo, H.-K., J. Yamamoto, T. Saito, and Y. Sakurai. 2014. Laboratory observations on the vertical swimming behavior of Japanese common squid *Todarodes pacificus* paralarvae as they ascend into warm surface waters. Fisheries Science 80:925–932.

York, C. A., and I. K. Bartol. 2016. Anti-predator behavior of squid throughout ontogeny. Journal of Experimental Marine Biology and Ecology 48':26–35.

Young, R. E., and R. F. Harman. 1985. Early life history stages of enoploteuthin squids (Cephalopoda : Teuthoidea : Enoploteuthidae) from Hawaiian waters. Vie et Milieu 35:181–201.

Young, R. E., and R. F. Harman. 1988. “Larva,” “paralarva” and “subadult” in cephalopod terminology. Malacologia 29:201–207.

Young, R. E., and K. M. Mangold. 1994. Growth and reproduction in the mesopelagic-boundary squid *Abralia trigonura*. Marine Biology 119:413–421.

Zhang, Y., C. Hu, D. J. Mcgillicuddy, Y. Liu, B. B. Barnes, V. H. Kourafalou, Y. Zhang, C. Hu, D. J. Liu, and Y. Kourafalou. 2024. Journal Pre-proof Mesoscale eddies in the Gulf of Mexico: A three-dimensional characterization based on global HYCOM Mesoscale eddies in the Gulf of Mexico: A three-dimensional characterization based on 1 global HYCOM.

Zuur, A. F., and K. C. J. Camphuysen. 2013. Generalized Additive Models applied on northern gannets. Pages 145–168 *in* A. F. Zuur, editor. A Beginner’s Guide to Generalized Additive Models with R. 2nd edition. Highland Statistics Ltd., Newburgh.

Zuur, A. F., M. Skern-Mauritzen, J. Aars, E. N. Ieno, and A. A. Saveliev. 2014. Additive mixed effects models applied on polar bear movement data. Pages 35–72 *in* A. F. Zuur, A. A. Saveliev, and E. N. Ieno, editors. A Beginner’s Guide to Generalized Additive Mixed Effect Models with R. Highland Statistics, Ltd., Newburg.

# Appendices

**Appendix 1. Bibliography of All R Packages Used in Study (Alphabetical Order by Package Name)**


ape; v5.7-1: Paradis E, Schliep K (2019). “ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R.” _Bioinformatics_, *35*, 526-528. doi:10.1093/bioinformatics/bty633 <https://doi.org/10.1093/bioinformatics/bty633>.

BiodiversityR; v2.15-4: Kindt R, Coe R (2005). _Tree diversity analysis. A manual and software for common statistical methods for ecological and biodiversity studies_. World Agroforestry Centre (ICRAF). ISBN 92-9059-179-X, <http://www.worldagroforestry.org/output/tree-diversity-analysis>.

car; v3.1-2: Fox J, Weisberg S (2019). _An R Companion to Applied Regression_, Third edition. Sage, Thousand Oaks CA. <https://socialsciences.mcmaster.ca/jfox/Books/Companion/>.

caret; v6.0-94: Kuhn, M. (2008). Building Predictive Models in R Using the caret Package. Journal of Statistical Software, 28(5), 1–26. https://doi.org/10.18637/jss.v028.i05

dplyr; v1.1.2: Wickham H, François R, Henry L, Müller K, Vaughan D (2023). _dplyr: A Grammar of Data Manipulation_. R package version 1.1.2, <https://CRAN.R-project.org/package=dplyr>.

geosphere; v1.5-18: Hijmans R (2022). _geosphere: Spherical Trigonometry_. R package version 1.5-18, <https://CRAN.R-project.org/package=geosphere>.

GGally; v2.1.2: Schloerke B, Cook D, Larmarange J, Briatte F, Marbach M, Thoen E, Elberg A, Crowley J (2021). _GGally: Extension to 'ggplot2'_. R package version 2.1.2, <https://CRAN.R-project.org/package=GGally>.

ggnewscale; v0.4.9: Campitelli E (2023). _ggnewscale: Multiple Fill and Colour Scales in 'ggplot2'_. R package version 0.4.9, <https://CRAN.R-project.org/package=ggnewscale>.

ggordiplots; v0.4.3: Quensen J, Simpson G, Oksanen J (2024). _ggordiplots: Make 'ggplot2' Versions of Vegan's Ordiplots_. R package version 0.4.3, <https://CRAN.R-project.org/package=ggordiplots>.

ggplot2; v3.4.4: H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

ggupbr; v0.6.0: Kassambara A (2023). _ggpubr: 'ggplot2' Based Publication Ready Plots_. R package version 0.6.0, <https://CRAN.R-project.org/package=ggpubr>.

ggrepel; v0.9.4: Slowikowski K (2023). _ggrepel: Automatically Position Non-Overlapping Text Labels with 'ggplot2'_. R package version 0.9.4, <https://CRAN.R-project.org/package=ggrepel>.

ggspatial; v1.1.9: Dunnington D (2023). _ggspatial: Spatial Data Framework for ggplot2_. R package version 1.1.9, <https://CRAN.R-project.org/package=ggspatial>.

ggtext; v0.1.2: Wilke C, Wiernik B (2022). _ggtext: Improved Text Rendering Support for 'ggplot2'_. R package version 0.1.2, <https://CRAN.R-project.org/package=ggtext>.

gridExtra; v2.3: Auguie B (2017). _gridExtra: Miscellaneous Functions for "Grid" Graphics_. R package version 2.3, <https://CRAN.R-project.org/package=gridExtra>.

Hmisc; v5.1-1: Harrell Jr F (2023). _Hmisc: Harrell Miscellaneous_. R package version 5.1-1, <https://CRAN.R-project.org/package=Hmisc>.

labdsv; v2.1-0: Roberts DW (2023). _labdsv: Ordination and Multivariate Analysis for Ecology_. R package version 2.1-0, <https://CRAN.R-project.org/package=labdsv>.

maps; v3.4.1: Becker RA, Wilks, AR, Brownrigg R, Minka TP, Deckmyn A (2023). _ maps: Draw Geographical Maps_. R package version 3.4.1, <https://CRAN.R-project.org/package=maps>.

mgcv; v1.8-42: Wood, S.N. (2017) Generalized Additive Models: An Introduction with R (2nd edition). Chapman and Hall/CRC.

mgcViz; v0.1.11: Fasiolo, M., Nedellec, R., Goude, Y. and Wood, S.N., 2020. Scalable visualization methods for modern generalized additive models. Journal of computational and Graphical Statistics, 29(1), pp.78-86.

readr; v2.1.4: Wickham H, Hester J, Bryan J (2023). _readr: Read Rectangular Text Data_. R package version 2.1.4, <https://CRAN.R-project.org/package=readr>.

readxl; v1.4.3: Wickham H, Bryan J (2023). _readxl: Read Excel Files_. R package version 1.4.3, <https://CRAN.R-project.org/package=readxl>.

rstatix; v0.7.2: Kassambara A (2023). _rstatix: Pipe-Friendly Framework for Basic Statistical Tests_. R package version 0.7.2, <https://CRAN.R-project.org/package=rstatix>.

RVAideMemoire; v0.9-83-3: Herve M (2023). _RVAideMemoire: Testing and Plotting Procedures for Biostatistics_. R package version 0.9-83-3, <https://CRAN.R-project.org/package=RVAideMemoire>.

| | |
|---|---|
| sf; v1.0-14: | Pebesma, E., & Bivand, R. (2023). Spatial Data Science: With Applications in R. Chapman and Hall/CRC. https://doi.org/10.1201/9780429459016 |
| suncalc; v0.5.1: | Thieurmel B, Elmarhraoui A (2022). _suncalc: Compute Sun Position, Sunlight Phases, Moon Position and Lunar Phase_. R package version 0.5.1, <https://CRAN.R-project.org/package=suncalc>. |
| terra; v1.7-55: | Hijmans R (2023). _terra: Spatial Data Analysis_. R package version 1.7-55, <https://CRAN.R-project.org/package=terra>. |
| vegan; v2.6-4: | Oksanen J, Simpson G, Blanchet F, Kindt R, Legendre P, Minchin P, O'Hara R, Solymos P, Stevens M, Szoecs E, Wagner H, Barbour M, Bedward M, Bolker B, Borcard D, Carvalho G, Chirico M, De Caceres M, Durand S, Evangelista H, FitzJohn R, Friendly M, Furneaux B, Hannigan G, Hill M, Lahti L, McGlinn D, Ouellette M, Ribeiro Cunha E, Smith T, Stier A, Ter Braak C, Weedon J (2022). _vegan: Community Ecology Package_. R package version 2.6-4, <https://CRAN.R-project.org/package=vegan>. |
| writexl; v1.4.2: | Ooms J (2023). _writexl: Export Data Frames to Excel 'xlsx' Format_. R package version 1.4.2, <https://CRAN.R-project.org/package=writexl>. |

## Appendix 2. Supplementary Methods and Results of Abundance Modeling

### A2.1. Methods

Steps of the methods not explicitly described in the main body of the thesis are described herein. Additionally, the individual steps of the model creation, selection, and validation process are described in more complete detail.

#### *A2.1.1. Data Exploration*

Cleveland dotplots were produced to identify any outliers in the continuous variables. Outliers were plotted versus spatial variables to determine if the collection location of the sample could explain the outlying values. Pair plots were used to identify any collinearity between variables. Correlation coefficients were produced for pairs of continuous variables, boxplots were produced for continuous-categorical pairs, scatterplots were produced for pairs of

continuous variables, and barplots were produced for pairs of categorical variables. Plots were examined to determine the presence and strength of collinearities between variables.

To objectively select uncorrelated variables for analysis, variance inflation factors were calculated. Specifically, squared scaled generalized variance inflation factors (ssGVIF) were calculated for all continuous and categorical variables (DeRuiter 2019). The variable with the highest ssGVIF was removed from the analysis and the ssGVIFs were recalculated. This process was repeated until all ssGVIFs were below a threshold value of 4.

Continuous variables with outliers were plotted against each other to see if the correlation between them was driven by the outlying samples. The correlation coefficient between those continuous variables was calculated without the outlying samples to confirm their effect. If the outlying values could not be explained by their location or other factors, suggesting they are due to CTD errors, and they had a large impact on correlations between variables, those samples were removed from the dataset.

All continuous variables were replotted to check for remaining outliers. The pairplot matrix was plotted with the subset of data to check for changes in collinearity after the removal of samples with outliers. Variable selection via variance inflation factors was repeated. The selection process ended when no more variables had ssGVIFs greater than 4 and all remaining variables were recorded for input into the model.

The percentage of samples that had a value of zero for the response variable was calculated for all samples to determine whether the data were zero-inflated. The percentage of zeroes was further calculated for each level of the categorical variables to determine whether the percentage varied between levels.

To visualize possible relationships between the response variable and the predictor variables, scatterplots and boxplots were created with the continuous and categorical variables, respectively, plotted against the response variable.

*A2.1.2. Model Creation*

As the data are counts, the modeling process began with a generalized additive model (GAM) with a Poisson distribution. All variables selected by the VIF process were included in the model, with the continuous variables as modified thin plate regression spline smoothers. The sampling effort variable, volume filtered, was included as the natural logarithm of the value. After the model was run, the dispersion statistic was calculated (Zuur and Camphuysen 2013). The model summary was produced and the deviance explained was recorded. The effective degrees of freedom (EDF) for each smoother and the estimate for the natural logarithm of volume filtered variable were checked.

Based on the model results, a modified model was built and run using the following changes. If the estimate for the natural logarithm of volume filtered variable was approximately equal to 1, it was included in the new model as an offset variable. If it was not approximately equal to 1, it was included in the new model as a smoother as the original value, not the natural logarithm. If any smoothers had an EDF approximately equal to 1, this indicated that the relationship was linear and the smoother was removed while the variable was retained in the model as a parametric term.

After the modified Poisson GAM was run, the dispersion statistic was calculated. The model summary was produced and the deviance explained was recorded. The EDF for each smoother was checked and if any were approximately equal to 1, the smoother was removed while the variable was retained in the model as a parametric term. If any variables were newly switched from smoothers to parametric terms, the secondarily modified model was run and the overdispersion was calculated. If no variables were newly switched from smoothers to parametric terms, the dispersion statistic of the originally modified model was examined. If the dispersion statistic of the model under consideration was greater than 1, the model was over-dispersed and a new model with a quasi-Poisson distribution was created.

The same process was used to create and analyze the quasi-Poisson GAM. If the dispersion statistic of the final modified model was still greater than 1, a new model with a negative binomial distribution was created. The same process as the Poisson and quasi-Poisson GAMs was used to create and analyze the negative binomial GAM, with the addition of extracting the θ value from the model output.

*A2.1.3. Model Selection*

After a model was produced with no smoothers with an EDF approximately equal to 1 and a dispersion statistic approximately equal to 1, a model selection process was carried out. For all hypothesis tests, a significance level of 5% was used. First, a multi-model analysis of variance (ANOVA) was used to compare the models with a dispersion statistic parameter approximately equal to 1. The Akaike information criterion (AIC) was also compared for the models. The model considered “better” by these two measures was selected for further model selection. First, all smoothers with an EDF approximately equal 0, which indicates that the term was shrunk to zero, were removed from the model formula. This model was run and the model summary checked to confirm that these variables were already not part of the model. Next, variables were removed using backwards selection based on p-value: the variable with the highest p-value in the model was removed and the model was re-run. The dispersion statistic was calculated and the model summary was produced.

The two models (one with the variable and one without) were compared using a multi-model ANOVA and the AIC. If either test indicated that the reduced model was better, the model summary of that model was examined and if any variables had p-values greater than 0.05, the variable with the new highest p-value was removed and the model was re-run. The process was continued as described above until no variables had insignificant p-values. At this point, the working model was considered the best model. The smooths functions and the partial effects of the parametric model components were plotted for interpretation.

*A2.1.4. Model Validation*

Model validation was conducted to verify whether the model determined to be the best met the assumptions of the technique used and could be used for a robust interpretation. The homogeneity of variance was verified by plotting the Pearson residuals versus the fitted values. Model misfit (or independence) was verified visually by plotting the Pearson residuals versus each covariate in the model and each covariate not in the model and a LOWESS (locally-weighted polynomial regression) smoother was added to each plot (continuous variables only) to determine that there was no relationship between the variable and the residuals.

To verify this mathematically, a linear model was created relating the Pearson residuals to the variables not in the model, one at a time. If the F-test on the linear model of the residuals versus the explanatory variable and the linear model of the residuals versus no variables was insignificant, then the absence of a relationship between the residuals and the variable was confirmed. If the F-test was significant, then the variable should have been included in the model. Normality was verified by creating a histogram and Q-Q plot of the Pearson residuals.

Influential observations were identified using Cook's distance. Three thresholds were used: absolute Cook's distance values greater than 0.5, absolute Cook's distance values greater than 4 divided by the number of samples, and absolute Cook's distance values greater than 3 times the mean Cook's distance value. The Cook's distance values were plotted with the thresholds as horizontal lines to visualize the samples that exceeded each threshold.

To determine whether there were site-specific patterns that were not captured in the model, a boxplot was created of the Pearson residuals plotted by site ID. Finally, independence between samples collected at multiple spatial locations was verified by calculating Moran's I using both the Pearson residuals and the original count data. Additionally, spatial correlation between samples collected at the same site versus at other sites was tested using Moran's I.

## A2.2. Results

Additional results not explicitly stated in the main body of the thesis are included here. Additionally, all models tested are listed. Model output and model equations are included for the selected best-fitting models. The same three distributions were used for modeling all three taxa. Their equations are specified as follows:

| | | | |
|---|---|---|---|
| Poisson: | $Y_i \sim \text{Poisson}(\mu_i)$ | $E(Y_i) = \mu_i$ | $\text{var}(Y_i) = \mu_i$ |
| Quasi-Poisson: | $Y_i \sim \text{Poisson}(\mu_i, \theta_i)$ | $E(Y_i) = \mu_i$ | $\text{var}(Y_i) = v_i(\mu_i) = \theta_i\mu_i$ |
| Negative binomial: | $Y_i \sim \text{NB}(\mu_i, \theta)$ | $E(Y_i) = \mu_i$ | $\text{var}(Y_i) = \mu_i + \mu^2_i/\theta$ |

### *A2.2.1. Data Exploration*

There were 11 salinity outliers with values of less than 33, while all remaining values fell between 34.5 and 36.7. These were all collected at 2 sites offshore of the Florida Panhandle, the closest to shore sites. These sample eleven samples also had outlier transmissivity values of 0%. Transmissivity values of 0% suggest an error with the CTD instrument. The remaining continuous variables did not have any outlying values.

The following variable pairs had high Pearson correlation coefficients: temperature and density, salinity and transmissivity, and salinity and density. There appeared to be strong correlation between temperature and year, temperature and season, and temperature and depth bin. Possible correlation was identified between fluorescence and depth bin, oxygen and depth bin, density and year, density and season, and density and depth bin. Five variables had ssGVIF values above the cut-off of 4: density, temperature, salinity, transmissivity, and season. The correlation between salinity and transmissivity was apparently driven by the eleven samples with outlying values for both variables – without those samples the correlation between them was -0.0001 as opposed to 0.871.

Due to the influence of these samples and because their values were potentially due to errors in the CTD instruments, the eleven samples were removed from the dataset, leaving 516

samples. Re-running the variance inflation factor function on the subset of data produced four variables with ssGVIFs above the cut-off: density, temperature, salinity, and season. Density was removed and the function re-run. Two variables had ssGVIFs above the cut-off: temperature and season. Temperature was removed and the function re-run. None of the remaining variables had a ssGVIF above the cut-off.

Sampling effort varied considerably in time and space (Tables A2.1 - A2.6). Almost equal numbers of samples were collected in 2009 and 2012, but less than a quarter of the number in 2011. More samples were collected in winter than in early fall and fewest were collected in spring. Many more samples were collected in the West Florida Shelf and GOM slope regions than in the other regions, especially the GOM basin region. More samples were collected in the two shallowest bins than in the four remaining bins.

Table A2.1. Sampling effort in time and space (single variable). (Based on 516 samples).

| *Variable* | | | | | | |
|---|---|---|---|---|---|---|
| *Year* | 2009 | 2011 | 2012 | | | |
| | 247 | 57 | 212 | | | |
| *Season* | Winter | Spring | Early Fall | | | |
| | 226 | 113 | 177 | | | |
| *Diel Period* | Day | Night | | | | |
| | 282 | 233 | | | | |
| *Region* | NW GOM Shelf | NC GOM Shelf | WFS | GOM Slope | GOM Basin | |
| | 86 | 56 | 144 | 191 | 39 | |
| *Depth Bin* | 0-20 m | 20-40 m | 40-60 m | 60-80 m | 80-100 m | 100-130 m |
| | 149 | 153 | 90 | 48 | 40 | 36 |

Table A2.2. Sampling effort by year and season. (Based on 516 samples.)

| | *Winter* | *Spring* | *Early Fall* |
|---|---|---|---|
| *2009* | 187 | 0 | 60 |
| *2011* | 0 | 0 | 57 |
| *2012* | 39 | 113 | 60 |

Table A2.3. Sampling effort by year and region. (Based on 516 samples).

| | *NW GOM Shelf* | *NC GOM Shelf* | *WFS* | *GOM Slope* | *GOM Basin* |
|---|---|---|---|---|---|
| *2009* | 56 | 32 | 55 | 98 | 6 |
| *2011* | 11 | 14 | 32 | 0 | 0 |
| *2012* | 19 | 10 | 57 | 93 | 33 |

Table A2.4. Sampling effort by season and region. (Based on 516 samples).

| | *NW GOM Shelf* | *NC GOM Shelf* | *WFS* | *GOM Slope* | *GOM Basin* |
|---|---|---|---|---|---|
| *Winter* | 30 | 25 | 59 | 106 | 6 |
| *Spring* | 0 | 0 | 15 | 65 | 33 |
| *Early Fall* | 56 | 31 | 70 | 20 | 0 |

Table A2.5. Sampling effort by year and depth bin. (Based on 516 samples).

| | *0-20 m* | *20-40 m* | *40-60 m* | *60-80 m* | *80-100 m* | *100-130 m* |
|---|---|---|---|---|---|---|
| *2009* | 57 | 62 | 41 | 33 | 27 | 27 |
| *2011* | 24 | 23 | 8 | 1 | 1 | 0 |
| *2012* | 68 | 68 | 41 | 14 | 12 | 9 |

Table A2.6. Sampling effort by season and depth bin. (Based on 516 samples).

| | *0-20 m* | *20-40 m* | *40-60 m* | *60-80 m* | *80-100 m* | *100-130 m* |
|---|---|---|---|---|---|---|
| *Winter* | 42 | 44 | 43 | 38 | 31 | 28 |
| *Spring* | 46 | 45 | 22 | 0 | 0 | 0 |
| *Early Fall* | 61 | 64 | 25 | 10 | 9 | 8 |

*A2.2.2. Enoploteuthidae*

There were no obvious relationships between count of Enoploteuthidae and the continuous predictor variables based on the scatterplots. Collinearity, described by the Pearson correlation coefficient, was very low between the count of Enoploteuthidae and the continuous variables. Additionally, there were no apparent relationships between count of Enoploteuthidae and the categorical predictor variables from a visual inspection of the boxplots.

From the entire set of 516 samples, 76% had a value of 0 from the response variable count of Enoploteuthidae. The percentage of zeroes was fairly consistent between years (70-80%), seasons (72-86%), cruises (70-86%), regions (67-85%), and, to a slightly lesser extent,

depth bins (63-92%). It did seem as though sites with at least one non-zero count were likely to have other samples with non-zero counts, suggesting there may be patterns due to repeated measurements from the same object (site) (Figure A2.1).

All models from the first portion of the model creation and selection process are given in Table A2.7. Models are listed sequentially in the order they were created. Changes between models in each distribution family include the conversion of the volume filtered term from a parametric to a smooth term and smooth terms being linear (edf ~= 1) and converted to parametric terms.

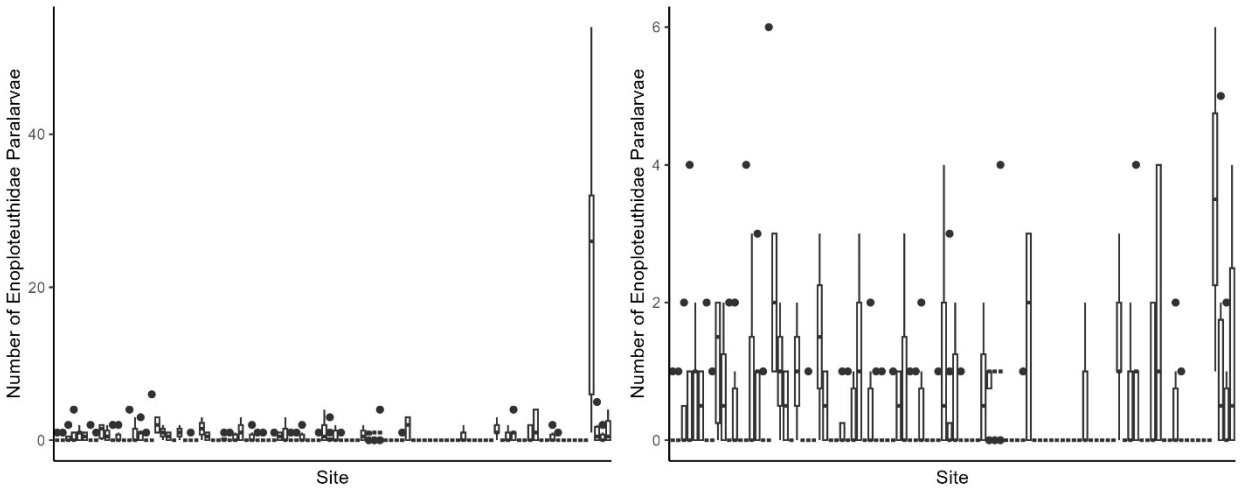


Figure A2.1. Boxplot of count of Enoploteuthidae for all samples (left) and samples with fewer than 10 paralarvae (right) by site showing potential correlation between abundance within sites.

Table A2.7. Model equations and summary statistics for Enoploteuthidae. In all equations μ is the count of Enoploteuthidae. (Quasi: quasi-Poisson; NB: negative binomial; Distrib.: distribution family; Dis.: dispersion statistic; Dev.: deviance explained; AR2: adjusted $R^2$; Est.: estimate for the parameter of natural logarithm of volume filtered)

| *Distrib.* | *Model Equation* | *Dis.* | *Dev.* | *AR2* | *Est.* | *θ* |
|---|---|---|---|---|---|---|
| *Poisson* | $\log(\mu_i) = \alpha + f_1(\text{fluorescence}_i) + f_2(\text{transmissity}_i) + f_3(\text{oxygen concentration}_i) + f_4(\text{salinity}_i) + \beta_1 \times \text{region}_i + \beta_2 \times \text{year}_i + \beta_3 \times \text{season}_i + \beta_4 \times \text{diel period}_i + \beta_5 \times \text{depth bin}_i + \beta_6 \times \ln(\text{volume filtered}_i)$ | 4.34 | 37.4% | 0.36 | 0.43998 | |
| *Poisson* | $\log(\mu_i) = \alpha + f_1(\text{fluorescence}_i) + f_2(\text{transmissity}_i) + f_3(\text{oxygen concentration}_i) + f_4(\text{salinity}_i) + f_5(\text{volume filtered}_i) + \beta_1 \times \text{region}_i + \beta_2 \times \text{year}_i + \beta_3 \times \text{season}_i + \beta_4 \times \text{diel period}_i + \beta_5 \times \text{depth bin}_i$ | 3.37 | 42.8% | 0.431 | | |
| *Quasi* | $\log(\mu_i) = \alpha + f_1(\text{fluorescence}_i) + f_2(\text{transmissity}_i) + f_3(\text{oxygen concentration}_i) + f_4(\text{salinity}_i) + \beta_1 \times \text{region}_i + \beta_2 \times \text{year}_i + \beta_3 \times \text{season}_i + \beta_4 \times \text{diel period}_i + \beta_5 \times \text{depth bin}_i + \beta_6 \times \ln(\text{volume filtered}_i)$ | 3.57 | 33% | 0.215 | 0.21860 | |
| *Quasi* | $\log(\mu_i) = \alpha + f_1(\text{fluorescence}_i) + f_2(\text{transmissity}_i) + f_3(\text{oxygen concentration}_i) + f_4(\text{salinity}_i) + )+ f_5(\text{volume filtered}_i) + \beta_1 \times \text{region}_i + \beta_2 \times \text{year}_i + \beta_3 \times \text{season}_i + \beta_4 \times \text{diel period}_i + \beta_5 \times \text{depth bin}_i$ | 3.10 | 38% | 0.278 | | |
| *NB* | $\log(\mu_i) = \alpha + f_1(\text{fluorescence}_i) + f_2(\text{transmissity}_i) + f_3(\text{oxygen concentration}_i) + f_4(\text{salinity}_i) + \beta_1 \times \text{region}_i + \beta_2 \times \text{year}_i + \beta_3 \times \text{season}_i + \beta_4 \times \text{diel period}_i + \beta_5 \times \text{depth bin}_i + \beta_6 \times \ln(\text{volume filtered}_i)$ | 1.09 | 22% | 0.0628 | 0.54519 | 0.2934 |
| *NB* | $\log(\mu_i) = \alpha + f_1(\text{fluorescence}_i) + f_2(\text{transmissity}_i) + f_3(\text{oxygen concentration}_i) + f_4(\text{salinity}_i) + f_5(\text{volume filtered}_i) + \beta_1 \times \text{region}_i + \beta_2 \times \text{year}_i + \beta_3 \times \text{season}_i + \beta_4 \times \text{diel period}_i + \beta_5 \times \text{depth bin}_i$ | 1.08 | 24.6% | 0.0767 | | 0.3111 |

Both the multi-model analysis of variance (ANOVA) and the AIC showed that the negative binomial model with volume filtered as a smoothed term was better (ANOVA p-value = 0.03772, AIC with smoothed term = 927.897 vs. AIC with parametric term = 931.832). The smoothers for fluorescence, transmissivity, and oxygen all had EDF approximately equal to zero, indicating they were selected out of the model during the model building process. A model was coded without them in order to compare with future reduced models.

For model selection, the region variable (p-value = 0.90) was removed and the model was re-run. The overdispersion is 1.179, indicating the model was likely not overdispersed. The deviance explained by the model is 21.4% and the adjusted $R^2$ is 0.0612. The estimated value of θ is 0.289. This model was compared with the model containing region. The multi-model ANOVA showed no difference between them, but the AIC of the model without region was lower (925.664 vs. 927.897). The EDF for the smoother for salinity was approximately equal to zero, indicating that it was selected out of the model during the model building process. A final GAM was created without salinity in the code, but the results of the model were not different than the previous.

The final model was as follows:

$count_i \sim NB(\mu_i, 0.289)$

$E(count_i) = \mu_i$

$var(count_i) = \mu_i + \mu^2_i/0.289 = \mu_i + 3.46 \times \mu^2_i$

$log(\mu_i) = -1.11 + f(\text{volume filtered}_i) + \beta_1 \times \text{year}_i + \beta_2 \times \text{season}_i + \beta_3 \times \text{diel period}_i + \beta_4 \times \text{depth bin}_i$

The effects of each term are plotted in Figure 7 in the main body of the thesis. The effects and significance of each parametric term as well as the model summary metrics are given in Table A2.8 and Table A2.9. The significance of each smooth term is given in Table A2.9. Model equations changing one variable at a time while the other variables are held at the base level are given in Table A2.10.

Table A2.8. Model summary output of the final Enoploteuthidae model. P-values of parametric terms are produced via ANOVA and describe the significance of the variable as a whole.

| **Parametric Terms** | | | | |
|---|---|---|---|---|
| | *df* | *Chi.sq* | *p-value* | |
| *Year* | 2 | 8.209 | 0.0165 | |
| *Season* | 2 | 20.077 | 4.37e-05 | |
| *Diel Period* | 1 | 3.856 | 0.0496 | |
| *Depth Bin* | 5 | 14.223 | 0.0143 | |
| **Approximate Significance of Smooth Terms** | | | | |
| | *edf* | *Ref.df* | *Chi.sq* | *p-value* |
| *Volume Filtered* | 3.362 | 9.000 | 11.36 | 0.00754 |
| **Model Metrics** | | | | |
| | Deviance Explained | Adjusted $R^2$ | θ | |
| | 21.4% | 0.0612 | 0.289 | |

Table A2.9. Parametric coefficients of the final Enoploteuthidae model. Estimates and significance are relative to the base level of the factor. Positive values of the intercept indicate that the count of paralarvae is higher relative to the count in the base level, while negative values indicate the opposite. Insignificant values mean that the level does not have significantly different counts than the base level.

| | *Estimate* | *Std. Error* | *z-value* | *p-value* |
|---|---|---|---|---|
| *Intercept* | -1.1100 | 0.3023 | -3.671 | **0.000241** |
| *Year, relative to 2009* | | | | |
| *2011* | -0.5214 | 0.4790 | -1.089 | 0.276325 |
| *2012* | 0.5758 | 0.3544 | 1.625 | 0.104258 |
| *Season, relative to Winter* | | | | |
| *Spring* | -1.1133 | 0.4134 | -2.693 | **0.007087** |
| *Early Fall* | 0.5544 | 0.3382 | 1.639 | 0.101135 |
| *Diel period, relative to Day* | | | | |
| *Night* | 0.4350 | 0.2215 | 1.964 | **0.049559** |
| *Depth bin, relative to 0-20 m* | | | | |
| *20-40 m* | 0.2993 | 0.2758 | 1.085 | 0.277926 |
| *40-60 m* | 0.1309 | 0.3231 | 0.405 | 0.685464 |
| *60-80 m* | -0.2087 | 0.4193 | -0.498 | 0.618571 |
| *80-100 m* | -1.0073 | 0.5121 | -1.967 | **0.049202** |
| *100-130 m* | -1.6248 | 0.6341 | -2.562 | **0.010401** |

Table A2.10. Model equations for each variable holding the other variables at the base level.

| | *Significance Relative to Base Level* | *Level* | *Model Equation* |
|---|---|---|---|
| *Year (season, diel period, and depth bin set to base level):* | | | |
| | NA | 2009 | $\log(\mu_i) = -1.1100 + f(\text{volume filtered}_i)$ |
| | Non-significant | 2011 | $\log(\mu_i) = -1.1100 + f(\text{volume filtered}_i) - 0.5214$ |
| | Non-significant | 2012 | $\log(\mu_i) = -1.1100 + f(\text{volume filtered}_i) + 0.5758$ |
| *Season (year, diel period, and depth bin set to base level):* | | | |
| | NA | Winter | $\log(\mu_i) = -1.1100 + f(\text{volume filtered}_i)$ |
| | Significant | Spring | $\log(\mu_i) = -1.1100 + f(\text{volume filtered}_i) - 1.1133$ |
| | Non-significant | Early Fall | $\log(\mu_i) = -1.11 + f(\text{volume filtered}_i) + 0.5544$ |
| *Diel period (year, season, and depth bin set to base level):* | | | |
| | NA | Day | $-1.1100 + f(\text{volume filtered}_i)$ |
| | Significant | Night | $-1.1100 + f(\text{volume filtered}_i) + 0.4350$ |
| *Depth bin (year, season, and diel period set to base level):* | | | |
| | NA | 0-20 m | $\log(\mu_i) = -1.1100 + f(\text{volume filtered}_i)$ |
| | Non-significant | 20-40 m | $\log(\mu_i) = -1.1100 + f(\text{volume filtered}_i) + 0.2993$ |
| | Non-significant | 40-60 m | $\log(\mu_i) = -1.1100 + f(\text{volume filtered}_i) + 0.1309$ |
| | Non-significant | 60-80 m | $\log(\mu_i) = -1.1100 + f(\text{volume filtered}_i) - 0.2087$ |
| | Significant | 80-100 m | $\log(\mu_i) = -1.1100 + f(\text{volume filtered}_i) - 1.0073$ |
| | Significant | 100-130 m | $\log(\mu_i) = -1.1100 + f(\text{volume filtered}_i) - 1.6248$ |

The model has perhaps some heterogeneity of variance (Figure A2.2). The Pearson residuals include much higher magnitude positive values than negative values (approximately 8 vs -1). Additionally, the positive values of the residuals decreased along the x-axis except for three high values, which were the three samples with the highest counts. This suggests the model was not very good at predicting very high values. Approximately one quarter of the residuals were positive while the remaining were 0 or negative. This suggests that the model more frequently overestimated values, although only by a small amount. It underestimated values less frequently, but by a larger amount.

There were no obvious patterns between the residuals and any of the variables in the model, indicating good model fit. Additionally, there were no obvious patterns between the residuals and any of the variables not in the model, indicating that none of the variables needed to be added into the model.

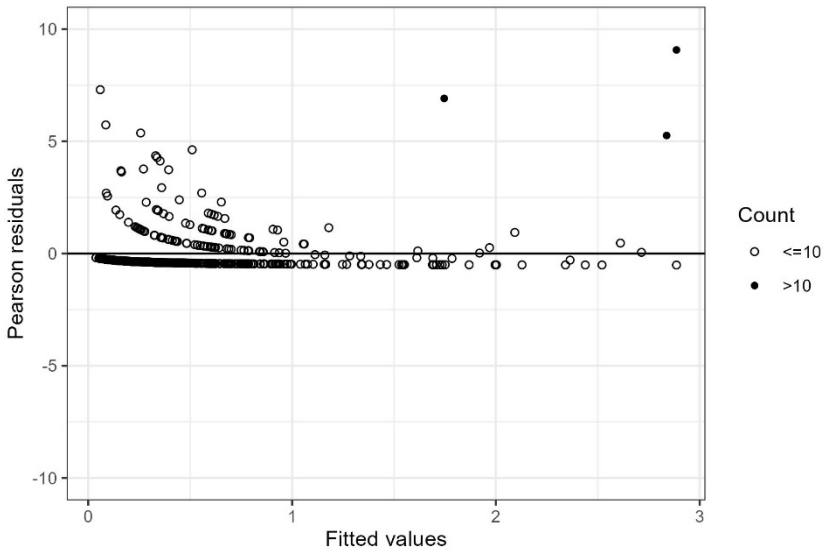


Figure A2.2. Plot of Pearson residuals versus fitted values indicating heterogeneity of variance. The filled circles indicate observations with large counts of Enoploteuthidae.

The histogram of the residuals is right-skewed, suggesting that the error terms are not normally distributed (Figure A2.3). The Q-Q plot indicates the same problem (Figure A2.3). The values mostly fall along the line except for at the right end of the plot, where they fall above the line. The Cook's distances suggest that there are no strongly influential points (Cook's distance greater than 0.5), but there are 23 potentially influential observations based on the threshold three times the mean of the Cook's distance.

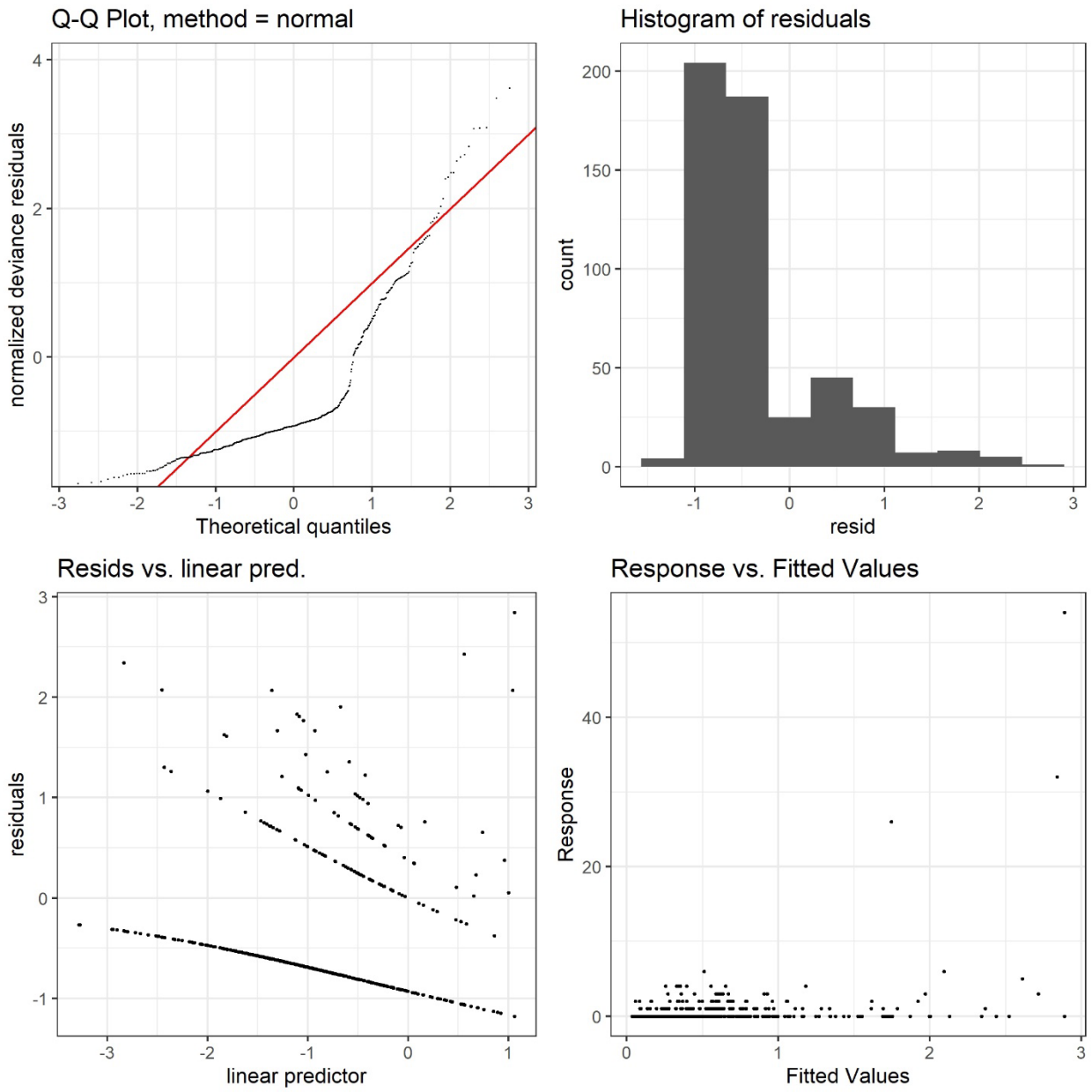


Figure A2.3. Model validation plots indicating violation of the assumptions of normality and heterogeneity of variance.

*A2.2.3. Octopodidae*

Based on the Pearson correlation coefficients and scatterplots, there were no continuous variables strongly related to count of Octopodidae. The boxplots also show no strong relationships between count of Octopodidae and the categorical variables.

From the entire set of 516 samples, 77.13% of the samples had a value of zero for the response variable, suggesting it may be zero-inflated. The percentage of samples with a count of Octopodidae equal to zero was variable between years (47-83%), seasons (66-93%), cruises (47-93%), regions (63-95%), and, to a lesser extent, depth bins (73-89%). There may be some correlation between samples taken at the same site based on the boxplot of site; it suggests that sites with one sample of above zero count are likely to have additional above-zero counts samples (Figure A2.4).

All models from the first portion of the model creation and selection process are given in Table A2.11. Models are listed sequentially in the order they were created. Changes between models in each distribution family include the conversion of the volume filtered term from a parametric to a smooth term and smooth terms being linear (edf ~= 1) and converted to parametric terms.

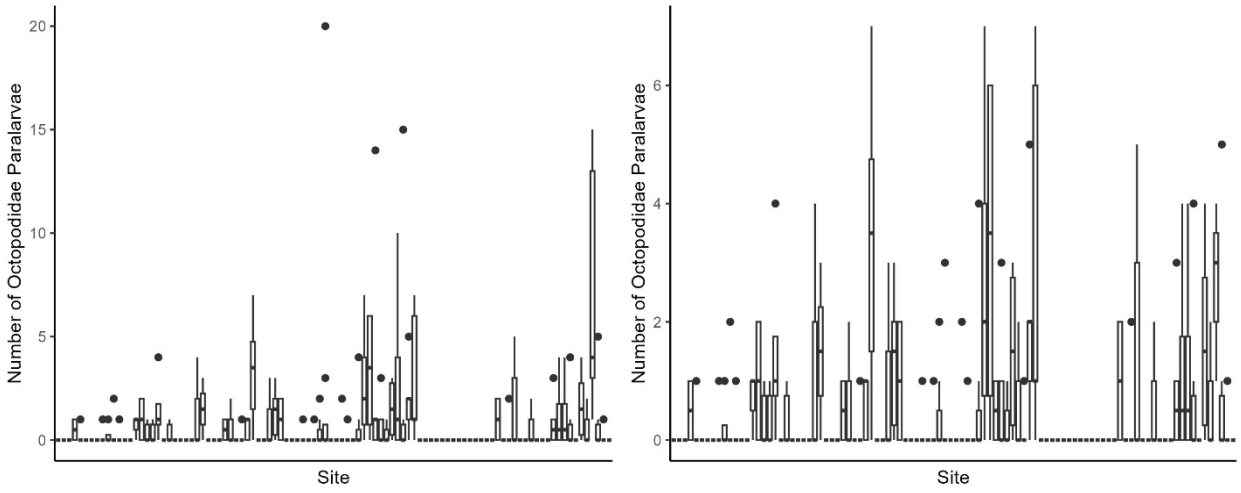


Figure A2.4. Boxplot of count of Octopodidae for all samples (left) and samples with fewer than 10 paralarvae (right) by site showing potential correlation between abundance within sites.

Table A2.11. Model equations and summary statistics for Octopodidae. In all equations μ is the count of Octopodidae. (Quasi: quasi-Poisson; NB: negative binomial; Distrib.: distribution family; Dis.: dispersion statistic; Dev.: deviance explained; AR2: adjusted $R^2$; Est.: estimate for the parameter of natural logarithm of volume filtered)

| *Distrib.* | *Model Equation* | *Dis.* | *Dev.* | *AR2* | *Est.* | *θ* |
|---|---|---|---|---|---|---|
| *Poisson* | $\log(\mu_i) = \alpha + f_1(\text{fluorescence}_i) + f_2(\text{transmissity}_i) + f_3(\text{oxygen concentration}_i) + f_4(\text{salinity}_i) + \beta_1 \times \text{region}_i + \beta_2 \times \text{year}_i + \beta_3 \times \text{season}_i + \beta_4 \times \text{diel period}_i + \beta_5 \times \text{depth bin}_i + \beta_6 \times \ln(\text{volume filtered}_i)$ | 3.198 | 34% | 0.213 | 0.66821 | |
| *Poisson* | $\log(\mu_i) = \alpha + f_1(\text{fluorescence}_i) + f_2(\text{transmissity}_i) + f_3(\text{oxygen concentration}_i) + f_4(\text{salinity}_i) + f_5(\text{volume filtered}_i) + \beta_1 \times \text{region}_i + \beta_2 \times \text{year}_i + \beta_3 \times \text{season}_i + \beta_4 \times \text{diel period}_i + \beta_5 \times \text{depth bin}_i$ | 3.231 | 35.4% | 0.243 | | |
| *Poisson* | $\log(\mu_i) = \alpha + f_1(\text{transmissity}_i) + f_2(\text{oxygen concentration}_i) + f_3(\text{volume filtered}_i) + \beta_1 \times \text{region}_i + \beta_2 \times \text{year}_i + \beta_3 \times \text{season}_i + \beta_4 \times \text{diel period}_i + \beta_5 \times \text{depth bin}_i + \beta_6 \times \text{fluorescence}_i + \beta_7 \times \text{salinity}_i$ | 3.249 | 35.3% | 0.243 | | |
| *Quasi* | $\log(\mu_i) = \alpha + f_1(\text{fluorescence}_i) + f_2(\text{transmissity}_i) + f_3(\text{oxygen concentration}_i) + f_4(\text{salinity}_i) + \beta_1 \times \text{region}_i + \beta_2 \times \text{year}_i + \beta_3 \times \text{season}_i + \beta_4 \times \text{diel period}_i + \beta_5 \times \text{depth bin}_i + \beta_6 \times \ln(\text{volume filtered}_i)$ | 3.686 | 29.4% | 0.176 | 0.77158 | |
| *Quasi* | $\log(\mu_i) = \alpha + f_1(\text{fluorescence}_i) + f_2(\text{transmissity}_i) + f_3(\text{oxygen concentration}_i) + f_4(\text{salinity}_i) + f_5(\text{volume filtered}_i) + \beta_1 \times \text{region}_i + \beta_2 \times \text{year}_i + \beta_3 \times \text{season}_i + \beta_4 \times \text{diel period}_i + \beta_5 \times \text{depth bin}_i$ | 3.592 | 29.5% | 0.173 | | |
| *Quasi* | $\log(\mu_i) = \alpha + f_1(\text{transmissity}_i) + \beta_1 \times \text{region}_i + \beta_2 \times \text{year}_i + \beta_3 \times \text{season}_i + \beta_4 \times \text{diel period}_i + \beta_5 \times \text{depth bin}_i + \beta_6 \times \text{fluorescence}_i + \beta_7 \times \text{salinity}_i + \beta_8 \times \text{oxygen concentration}_i + \beta_9 \times \text{volume filtered}_i$ | 3.625 | 29.5% | 0.171 | | |
| *NB* | $\log(\mu_i) = \alpha + f_1(\text{fluorescence}_i) + f_2(\text{transmissity}_i) + f_3(\text{oxygen concentration}_i) + f_4(\text{salinity}_i) + \beta_1 \times \text{region}_i + \beta_2 \times \text{year}_i + \beta_3 \times \text{season}_i + \beta_4 \times \text{diel period}_i + \beta_5 \times \text{depth bin}_i + \beta_6 \times \ln(\text{volume filtered}_i)$ | 1.255 | 25.7% | 0.141 | 0.83442 | 0.298 |
| *NB* | $\log(\mu_i) = \alpha + f_1(\text{fluorescence}_i) + f_2(\text{transmissity}_i) + f_3(\text{oxygen concentration}_i) + f_4(\text{salinity}_i) + f_5(\text{volume filtered}_i) + \beta_1 \times \text{region}_i + \beta_2 \times \text{year}_i + \beta_3 \times \text{season}_i + \beta_4 \times \text{diel period}_i + \beta_5 \times \text{depth bin}_i$ | 1.231 | 25.9% | 0.141 | | 0.298 |
| *NB* | $\log(\mu_i) = \alpha + f_1(\text{transmissity}_i) + f_2(\text{oxygen concentration}_i) + \beta_1 \times \text{region}_i + \beta_2 \times \text{year}_i + \beta_3 \times \text{season}_i + \beta_4 \times \text{diel period}_i + \beta_5 \times \text{depth bin}_i + \beta_6 \times \text{fluorescence}_i + \beta_7 \times \text{salinity}_i + \beta_8 \times \text{volume filtered}_i$ | 1.275 | 26.1% | 0.141 | | 0.306 |

Multi-model ANOVA failed to reject the null hypothesis that the models explained the same amount of deviance, indicating that none of the negative binomial models was any better than the others. The AIC also indicated no difference between the models, as the ΔAIC was less than 2 for all of the models (NB1: 924.0747, NB2: 923.3154; NB3: 922.8681). Since none of the models were considered significantly better, the third model was chosen for further model selection. Since the two smoothers in this model, transmissivity and oxygen concentration, had EDF approximately equal to zero, they were already selected out of the model. The model was re-coded to remove them from the model formula. This essentially converted the model to a generalized linear model, as no smoothers were included. A backwards selection process was carried out on this model to find the best-fitting model (Table A2.12).

Table A2.12. Models produced by the backwards selection process. The first row is the model chosen by the selection process above. The dispersion statistic, deviance explained, adjusted $R^2$, θ, and AIC are for the model produced after the variable listed in the first column was removed. The multi-model ANOVA column indicates whether the model produced after the variable was removed is significantly different than the model with that variable. (ns: non-significant)

| *Variable Removed* | *p-value* | *Dispersion Statistic* | *Deviance Explained* | *Adjusted $R^2$* | *θ* | *Multi-model ANOVA* | *AIC* |
|---|---|---|---|---|---|---|---|
| *None* | NA | 1.27467 | 26.1% | 0.141 | 0.306 | NA | 922.8675 |
| *Diel Period* | 0.977 | 1.272 | 26.1% | 0.143 | 0.306 | ns | 920.8682 |
| *Year* | 0.792 | 1.243 | 26% | 0.146 | 0.305 | ns | 917.2679 |
| *Depth Bin* | 0.24 | 1.236 | 24.3% | 0.129 | 0.292 | ns | 913.6607 |
| *Salinity* | 0.149 | 1.207 | 23.8% | 0.122 | 0.287 | ns | 913.4876 |

The final negative binomial model was specified as follows:

$count_i \sim NB(\mu_i, 0.287)$

$E(count_i) = \mu_i$

$var(count_i) = \mu_i + \mu^2_i/0.287 = \mu_i + 3.48 \times \mu^2_i$

$log(\mu_i) = \alpha + \beta_1 \times region_i + \beta_2 \times season_i + \beta_3 \times fluorescence_i + \beta_4 \times volume\ filtered_i$

The effects of each term are plotted in Figure 8 in the main body of the thesis. The effects and significance of each parametric term as well as the model summary metrics are given in Table A2.13 and Table A2.14. Model equations changing one variable at a time while the other variables are held at the base level are given in Table A2.15.

Table A2.13. Model summary. P-values of parametric terms are produced via ANOVA and describe the significance of the variable as a whole.

| **Parametric Terms** | | | |
|---|---|---|---|
| | *df* | *Chi.sq* | *p-value* |
| *Region* | 4 | 18.636 | 9.26e-04 |
| *Season* | 2 | 21.038 | 2.7e-05 |
| *Fluorescence* | 1 | 13.390 | 2.53e-04 |
| *Volume Filtered* | 1 | 4.842 | 0.0278 |
| **Model Metrics** | | | |
| | Deviance Explained | Adjusted $R^2$ | θ |
| | 23.8% | 0.122 | 0.287 |

Table A2.14. Parametric coefficients of the final Octopodidae model. Estimates and significance are relative to the base level of the factor for categorical variables. Positive values of the intercept indicate that the count of paralarvae is higher relative to the count in the base level, while negative values indicate the opposite. Insignificant values mean that the level does not have significantly different counts than the base level. Significant values are bolded.

| | *Estimate* | *Std. Error* | *z-value* | *p-value* |
|---|---|---|---|---|
| *Intercept* | -1.9726 | 0.4539 | -4.346 | **1.39e-05** |
| *Fluorescence* | 0.3902 | 0.1066 | 3.659 | **0.000253** |
| *Volume Filtered* | 0.0030 | 0.0014 | 2.201 | **0.027769** |
| *Region, relative to NW GOM Shelf* | | | | |
| *NC GOM Shelf* | -1.0597 | 0.4420 | -2.398 | **0.016499** |
| *WFS* | 0.0326 | 0.3243 | 0.100 | 0.919954 |
| *GOM Slope* | -0.9051 | 0.3616 | -2.503 | **0.012305** |
| *GOM Basin* | -1.6902 | 0.7577 | -2.231 | **0.025702** |
| *Season, relative to Winter* | | | | |
| *Spring* | -0.0207 | 0.3935 | -0.053 | 0.958129 |
| *Early Fall* | 1.1479 | 0.2594 | 4.425 | **9.66e-06** |

Table A2.15. Model equations for each categorical variable holding the other variable at the base level.

| | *Significance Relative to Base Level* | *Level* | *Model Equation* |
|---|---|---|---|
| *Region (season set to base level):* | | | |
| | NA | NW GOM Shelf | $\log(\mu_i) = -1.973 + 0.390 \times \text{fluorescence}_i + 0.003 \times \text{volume filtered}_i$ |
| | Significant | NC GOM Shelf | $\log(\mu_i) = -1.973 - 1.060 + 0.390 \times \text{fluorescence}_i + 0.003 \times \text{volume filtered}_i$ |
| | Non-significant | WFS | $\log(\mu_i) = -1.973 + 0.033 + 0.390 \times \text{fluorescence}_i + 0.003 \times \text{volume filtered}_i$ |
| | Significant | GOM Slope | $\log(\mu_i) = -1.973 - 0.905 + 0.390 \times \text{fluorescence}_i + 0.003 \times \text{volume filtered}_i$ |
| | Significant | GOM Basin | $\log(\mu_i) = -1.973 - 1.690 + 0.390 \times \text{fluorescence}_i + 0.003 \times \text{volume filtered}_i$ |
| *Season (region set to base level):* | | | |
| | NA | Winter | $\log(\mu_i) = -1.973 + 0.390 \times \text{fluorescence}_i + 0.003 \times \text{volume filtered}_i$ |
| | Non-significant | Spring | $\log(\mu_i) = -1.973 - 0.021 + 0.390 \times \text{fluorescence}_i + 0.003 \times \text{volume}$ |
| | Significant | Early Fall | $\log(\mu_i) = -1.973 + 1.148 + 0.390 \times \text{fluorescence}_i + 0.003 \times \text{volume filtered}_i$ |

The model likely has slight heterogeneity of variance. The scatterplot of Pearson residuals versus fitted values shows that most residuals are small negative values, meaning that the model more frequently overestimated values, although only by a small amount (Figure A2.5). The positive residuals have higher magnitudes, indicating that the model underestimated values less frequently, but by a larger amount. The Pearson residuals include much higher magnitude positive values than negative values (approximately 8 vs -1). Additionally, the positive values of the residuals decreased along the x-axis. The samples with the highest counts of Octopodidae had generally high residuals, although only one of the three highest residuals was from a sample with a high count. This suggests that the model may not perform well in fitting the observations with high counts.

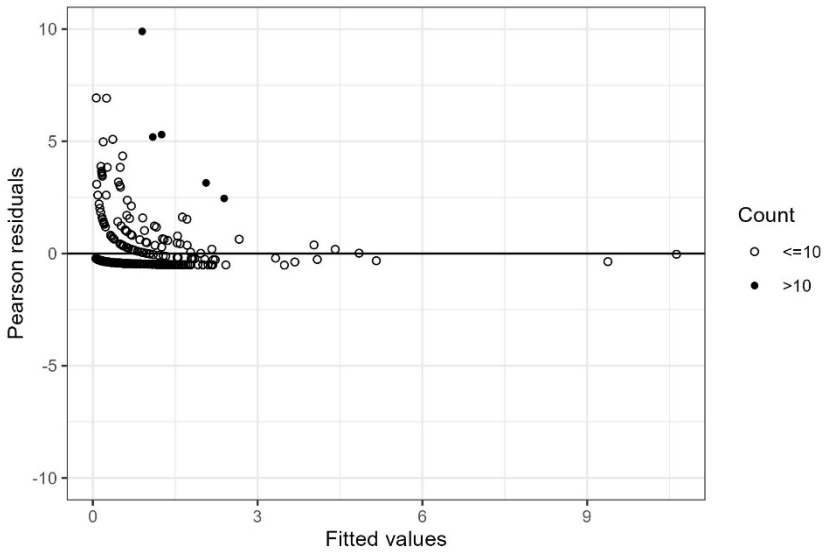


Figure A2.5. Plot of Pearson residuals versus fitted values indicating heterogeneity of variance. The filled circles indicate observations with large counts of Octopodidae.

The scatterplots and boxplots of residuals versus covariates in and not in the model do not show any patterns, indicating that the model fits well and that the samples are independent. This was confirmed by conducting F-tests on linear models of the residuals versus each variable not in the model individually. The histogram of the residuals is right-skewed, suggesting that the error terms are not normally distributed (Figure A2.6). The qqplot indicates the same problem (Figure A2.6). The values mostly fall along the line except for at the right end of the plot, where they fall above the line. The Cook's distances suggest that there are no strongly influential points (Cook's distance greater than 0.5), but there are 26 potentially influential observations based on the threshold three times the mean of the Cook's distance. The boxplot of the residuals versus site suggests that samples within sites may be correlated – residuals at a single site appeared more likely to be similar than residuals from different sites.

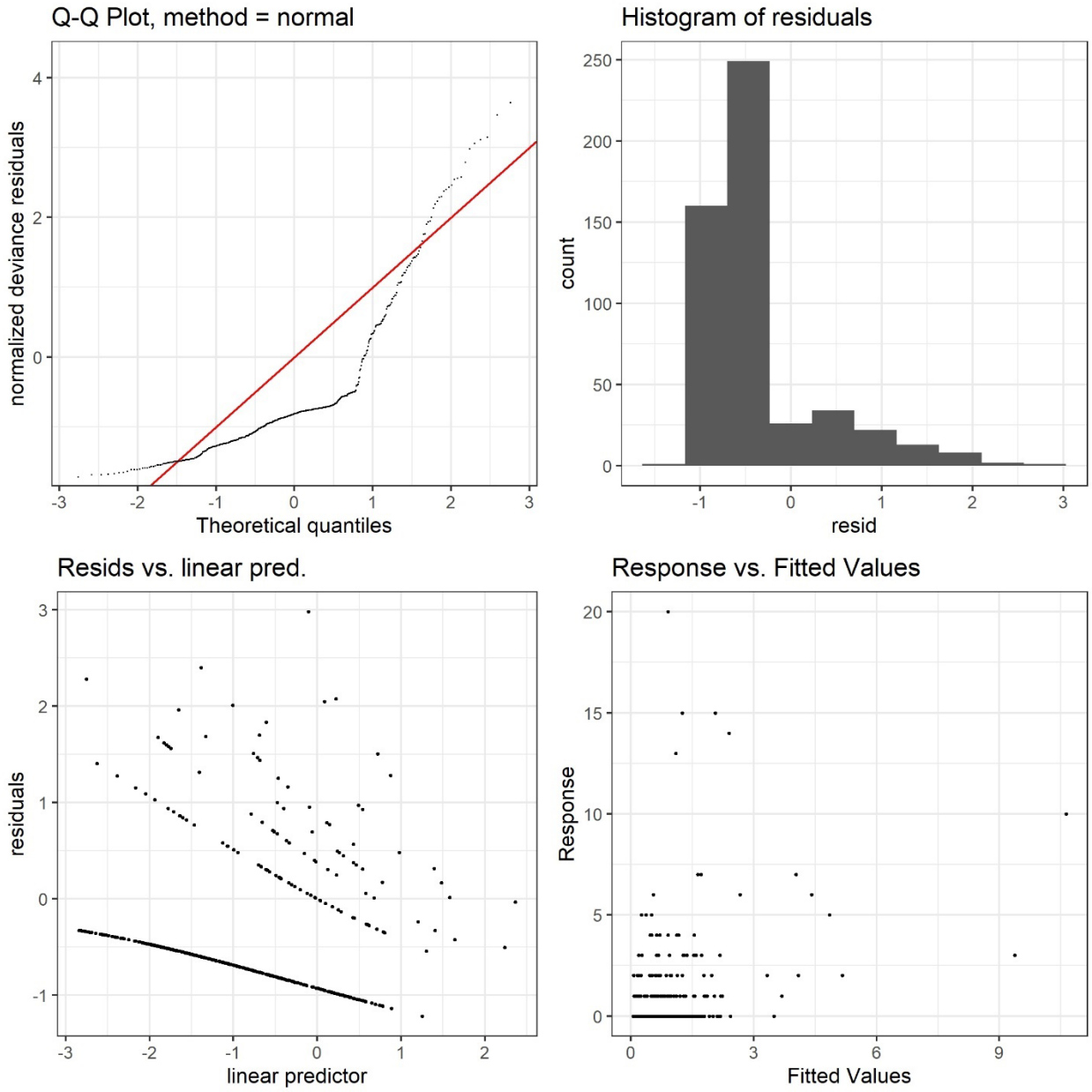


Figure A2.6. Model validation plots indicating violation of the assumptions of normality and heterogeneity of variance.

*A2.2.4. Loliginidae*

Based on the Pearson correlation coefficients and scatterplots, there were no continuous variables strongly related to count of Loliginidae. The boxplots also showed no strong relationships between count of Loliginidae and the categorical variables.

From the entire set of 516 samples, 92.83% had a value of 0 from the response variable count of Loliginidae. The percentage of samples with a count of Loliginidae equal to zero was variable between years (72-98%) and reached 100% for at least one level of the remaining categorical variables: seasons (85-100%), cruises (72-100%), regions (79-100%), and depth bins (90-100%). Based on the boxplot of count of Loliginidae versus site, it appears that most sites had no Loliginids in any samples, but those that did had multiple samples with Loliginids (Figure A2.7). This suggests there may be correlation between samples at a single site.

All models from the first portion of the model creation and selection process are given in Table A2.16. Models are listed sequentially in the order they were created. Changes between models in each distribution family include the conversion of the volume filtered term from a parametric to a smooth term and smooth terms being linear (edf ~= 1) and converted to parametric terms.

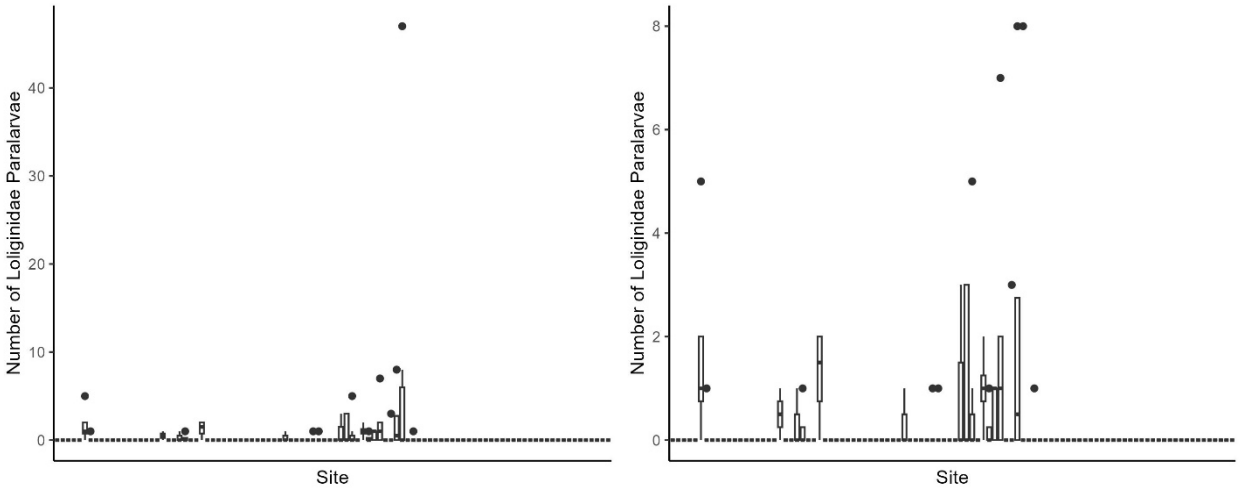


Figure A2.7. Boxplot of count of Loliginidae for all samples (left) and samples with fewer than 10 paralarvae (right) by site showing potential correlation between abundance within sites.

Table A2.16. Model equations and summary statistics for Loliginidae. In all equations μ is the count of Loliginidae. (Quasi: quasi-Poisson; NB: negative binomial; Distrib.: distribution family; Dis.: dispersion statistic; Dev.: deviance explained; AR2: adjusted $R^2$; Est.: estimate for the parameter of natural logarithm of volume filtered)

| *Distrib.* | *Model Equation* | *Dis.* | *Dev.* | *AR2* | *Est.* | *θ* |
|---|---|---|---|---|---|---|
| *Poisson* | $\log(\mu_i) = \alpha + f_1(\text{fluorescence}_i) + f_2(\text{transmissity}_i) + f_3(\text{oxygen concentration}_i) + f_4(\text{salinity}_i) + \beta_1 \text{ x region}_i + \beta_2 \text{ x year}_i + \beta_3 \text{ x season}_i + \beta_4 \text{ x diel period}_i + \beta_5 \text{ x depth bin}_i + \beta_6 \text{ x ln(volume filtered}_i)$ | 2.90 | 64.5% | 0.688 | -1.449 | |
| *Poisson* | $\log(\mu_i) = \alpha + f_1(\text{fluorescence}_i) + f_2(\text{transmissity}_i) + f_3(\text{oxygen concentration}_i) + f_4(\text{salinity}_i) + f_5(\text{volume filtered}_i) + \beta_1 \text{ x region}_i + \beta_2 \text{ x year}_i + \beta_3 \text{ x season}_i + \beta_4 \text{ x diel period}_i + \beta_5 \text{ x depth bin}_i$ | 2.61 | 64.8% | 0.702 | | |
| *Poisson* | $\log(\mu_i) = \alpha + f_1(\text{fluorescence}_i) + f_2(\text{transmissity}_i) + f_3(\text{oxygen concentration}_i) + f_4(\text{salinity}_i) + \beta_1 \text{ x region}_i + \beta_2 \text{ x year}_i + \beta_3 \text{ x season}_i + \beta_4 \text{ x diel period}_i + \beta_5 \text{ x depth bin}_i + \beta_6 \text{ x volume filtered}_i$ | 2.59 | 64.8% | 0.706 | | |
| *Quasi* | $\log(\mu_i) = \alpha + f_1(\text{fluorescence}_i) + f_2(\text{transmissity}_i) + f_3(\text{oxygen concentration}_i) + f_4(\text{salinity}_i) + \beta_1 \text{ x region}_i + \beta_2 \text{ x year}_i + \beta_3 \text{ x season}_i + \beta_4 \text{ x diel period}_i + \beta_5 \text{ x depth bin}_i + \beta_6 \text{ x ln(volume filtered}_i)$ | 2.48 | 66.4% | 0.743 | -1.204 | |
| *Quasi* | $\log(\mu_i) = \alpha + f_1(\text{fluorescence}_i) + f_2(\text{transmissity}_i) + f_3(\text{oxygen concentration}_i) + f_4(\text{salinity}_i) + f_5(\text{volume filtered}_i) + \beta_1 \text{ x region}_i + \beta_2 \text{ x year}_i + \beta_3 \text{ x season}_i + \beta_4 \text{ x diel period}_i + \beta_5 \text{ x depth bin}_i$ | 2.61 | 65.2% | 0.724 | | |
| *Quasi* | $\log(\mu_i) = \alpha + f_1(\text{fluorescence}_i) + f_2(\text{transmissity}_i) + f_3(\text{oxygen concentration}_i) + f_4(\text{salinity}_i) + \beta_1 \text{ x region}_i + \beta_2 \text{ x year}_i + \beta_3 \text{ x season}_i + \beta_4 \text{ x diel period}_i + \beta_5 \text{ x depth bin}_i + \beta_6 \text{ x volume filtered}_i$ | 2.33 | 66.4% | 0.749 | | |
| *NB* | $\log(\mu_i) = \alpha + f_1(\text{fluorescence}_i) + f_2(\text{transmissity}_i) + f_3(\text{oxygen concentration}_i) + f_4(\text{salinity}_i) + \beta_1 \text{ x region}_i + \beta_2 \text{ x year}_i + \beta_3 \text{ x season}_i + \beta_4 \text{ x diel period}_i + \beta_5 \text{ x depth bin}_i + \beta_6 \text{ x ln(volume filtered}_i)$ | 0.806 | 65.2% | -0.0535 | -1.54 | 0.233 |
| *NB* | $\log(\mu_i) = \alpha + f_1(\text{fluorescence}_i) + f_2(\text{transmissity}_i) + f_3(\text{oxygen concentration}_i) + f_4(\text{salinity}_i) + f_5(\text{volume filtered}_i) + \beta_1 \text{ x region}_i + \beta_2 \text{ x year}_i + \beta_3 \text{ x season}_i + \beta_4 \text{ x diel period}_i + \beta_5 \text{ x depth bin}_i$ | 0.838 | 65.3% | -0.0665 | | 0.221 |
| *NB* | $\log(\mu_i) = \alpha + f_1(\text{transmissity}_i) + f_2(\text{oxygen concentration}_i) + f_3(\text{salinity}_i) + \beta_1 \text{ x region}_i + \beta_2 \text{ x year}_i + \beta_3 \text{ x season}_i + \beta_4 \text{ x diel period}_i + \beta_5 \text{ x depth bin}_i + \beta_6 \text{ x fluorescence}_i + \beta_7 \text{ x volume filtered}_i$ | 0.754 | 65.2% | -0.129 | | 0.223 |

Multi-model ANOVA and AIC both indicated there were no differences between the three negative binomial models. Further model selection only produced underdispersed models with negative adjusted $R^2$ values, suggesting that the negative binomial distribution is not appropriate for this data. The three quasi-Poisson models, which were all overdispersed, had very similar deviance explained and adjusted $R^2$ values, suggesting that there were no differences between the three models. Since none of the models were considered significantly better, the third model was chosen for further model selection. Since two smoothers in this model, transmissivity and salinity, had EDF approximately equal to zero, they were already selected out of the model. The model was re-coded to remove them from the model formula. A backwards selection process was carried out on this model to find the best-fitting model (Table A2.17).

Table A2.17. Models produced by the backwards selection process. The first row is the model chosen by the selection process above. The dispersion statistic, deviance explained, and adjusted $R^2$ are for the model produced after the variable listed in the first column was removed.

| *Variable Removed* | *p-value* | *Dispersion Statistic* | *Deviance Explained* | *Adjusted $R^2$* |
|---:|---|---|---|---|
| *None* | NA | 2.328 | 66.4% | 0.749 |
| *Region* | 1 | 2.574 | 64.4% | 0.728 |
| *Season* | 1 | 2.131 | 68.2% | 0.774 |
| *Depth Bin* | 1 | 1.718 | 66.1% | 0.748 |
| *Oxygen Concentration* | 0.672 | 4.434 | 56.3% | 0.440 |
| *Salinity* | 0.995 | 4.434 | 56.3% | 0.440 |
| *Diel Period* | 0.294 | 1.780 | 66.0% | 0.750 |

As the model is still overdispersed, the quasi-Poisson distribution may not be appropriate for these data. Therefore, the results of the final model should be interpreted with caution. The final quasi-Poisson model was specified as follows:

$count_i \sim Poisson(\mu_i, \theta_i)$

$E(count_i) = \mu_i$

$var(count_i) = v_i(\mu_i) = \theta_i\mu_i$

$log(\mu_i) = \alpha + f_1(fluorescence_i) + f_2(transmissity_i) + \beta_1 \times year_i + \beta_2 \times volume\ filtered_i$

The effects of each term are plotted in Figure 9 in the main body of the thesis. The effects and significance of each parametric term as well as the model summary metrics are given in Table A2.18 and Table A2.19. The significance of each smooth term is given in Table A2.19. Model equations showing each level of the categorical variable are given in Table A2.20.

Table A2.18. Model summary. P-values of parametric terms are produced via ANOVA and describe the significance of the variable as a whole.

| **Parametric Terms** | | | | |
|---|---|---|---|---|
| | *df* | *F* | *p-value* | |
| *Year* | 2 | 30.92 | 2.21e-13 | |
| *Volume Filtered* | 1 | 21.24 | 5.16e-06 | |
| **Approximate Significance of Smooth Terms** | | | | |
| | *edf* | *Ref.df* | *F* | *p-value* |
| *Transmissivity* | 7.364 | 9.000 | 4.014 | 5.77e-06 |
| *Fluorescence* | 7.352 | 9.000 | 8.364 | < 2e-16 |
| **Model Metrics** | | | | |
| | Deviance Explained | Adjusted $R^2$ | | |
| | 66% | 0.75 | | |

Table A2.19. Parametric coefficients of the final Loliginidae model. Estimates and significance are relative to the base level of the factor. Positive values of the intercept indicate that the count of paralarvae is higher relative to the count in the base level, while negative values indicate the opposite. Insignificant values mean that the level does not have significantly different counts than the base level. Significant p-values are bolded.

| | *Estimate* | *Std. Error* | *t-value* | *p-value* |
|---|---|---|---|---|
| *Intercept* | -2.1581 | 0.8029 | -2.688 | **0.00743** |
| *Volume Filtered* | -0.0111 | 0.0024 | -4.609 | **5.16e-06** |
| *Year, relative to 2009* | | | | |
| *2011* | 3.1644 | 0.4068 | 7.780 | **4.23e-14** |
| *2012* | 1.7983 | 0.7348 | 2.447 | **0.01474** |

Table A2.20. Model equations showing the effect of each level of the categorical variable year.

| | *Significance Relative to Base Level* | *Level* | *Model Equation* |
|---|---|---|---|
| *Year:* | | | |
| | NA | 2009 | $\log(\mu_i) = -2.158 + f(\text{fluorescence}_i) + f(\text{transmissity}_i) - 0.011 \times \text{volume filtered}_i$ |
| | Significant | 2011 | $\log(\mu_i) = -2.158 + 3.164 + f(\text{fluorescence}_i) + f(\text{transmissity}_i) - 0.011 \times \text{volume filtered}_i$ |
| | Significant | 2012 | $\log(\mu_i) = -2.158 + 1.798 + f(\text{fluorescence}_i) + f(\text{transmissity}_i) - 0.011 \times \text{volume filtered}_i$ |

Model validation shows heterogeneity of variance, indicating a violation of the assumption of homogeneity of variance (Figure A2.8). The scatterplot of Pearson residuals versus fitted values shows that most residuals are small negative values, meaning that the model more frequently overestimated values, although only by a small amount. The positive residuals have higher magnitudes, indicating that the model underestimated values less frequently, but by a larger amount. The Pearson residuals include much higher magnitude positive values than negative values (approximately 12 vs. -3). Additionally, the positive values of the residuals decreased along the x-axis, except for the residual with the highest fitted value, which had a medium residual value. There was a large gap in fitted values between 8 and a single point with a fitted value of approximately 30. The sample with the outlying value for count of Loliginidae had a fairly high residual value, but other samples had higher residuals. This suggests that the model may not perform well in fitting the observations with very high counts, but possibly also not observations with medium counts.

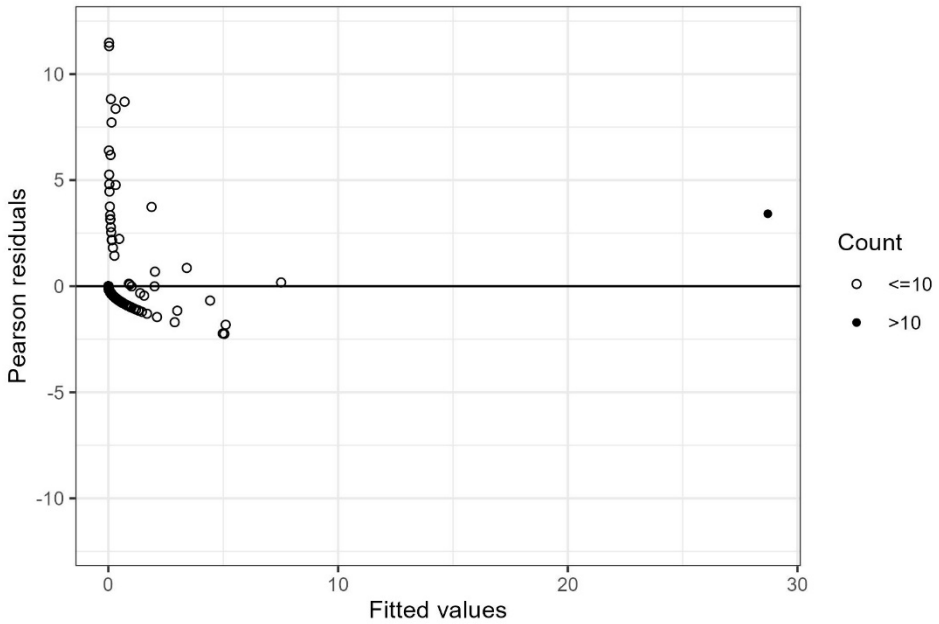


Figure A2.8. Plot of Pearson residuals versus fitted values indicating heterogeneity of variance. The filled circles indicate observations with large counts of Loliginidae.

The scatterplots of residuals versus covariates in the model appear to show slight patterns, but follow-up testing with F-tests on linear models of the residuals versus each variable individually did not find any significant relationships between the residuals and the variables. The boxplot of residuals versus year did not show any patterns. The scatterplots and boxplots of residuals versus covariates not in the model do not show any patterns, indicating that the model fits well and that the samples are independent. This was confirmed for all variables not in the model using F-tests as described above, except for the variable diel period. The F-test for this variable was significant, indicating that there is a relationship between the residuals and diel period, which suggests that it should have been included in the model.

The histogram of the residuals Is right-skewed, suggesting that the error terms are not normally distributed (Figure A2.9). Most samples fall into a single bin (-1 to 0) and very few samples fall in either tail. The Q-Q plot indicates the same problem (Figure A2.9). The values mostly fall along the line in the middle portion of the plot, but they fall slightly below the line at the left side and extend far above the line at the right side.

The Cook's distances suggest that there are two strongly influential points (Cook's distance greater than 0.5), which have a noticeable effect on the model produced on the set of samples containing them. Additionally, there are ten potentially influential observations based on the threshold three times the mean of the Cook's distance. The boxplot of the residuals versus site suggests that samples within sites may be correlated – residuals at a single site appeared more likely to be similar than residuals from different sites.

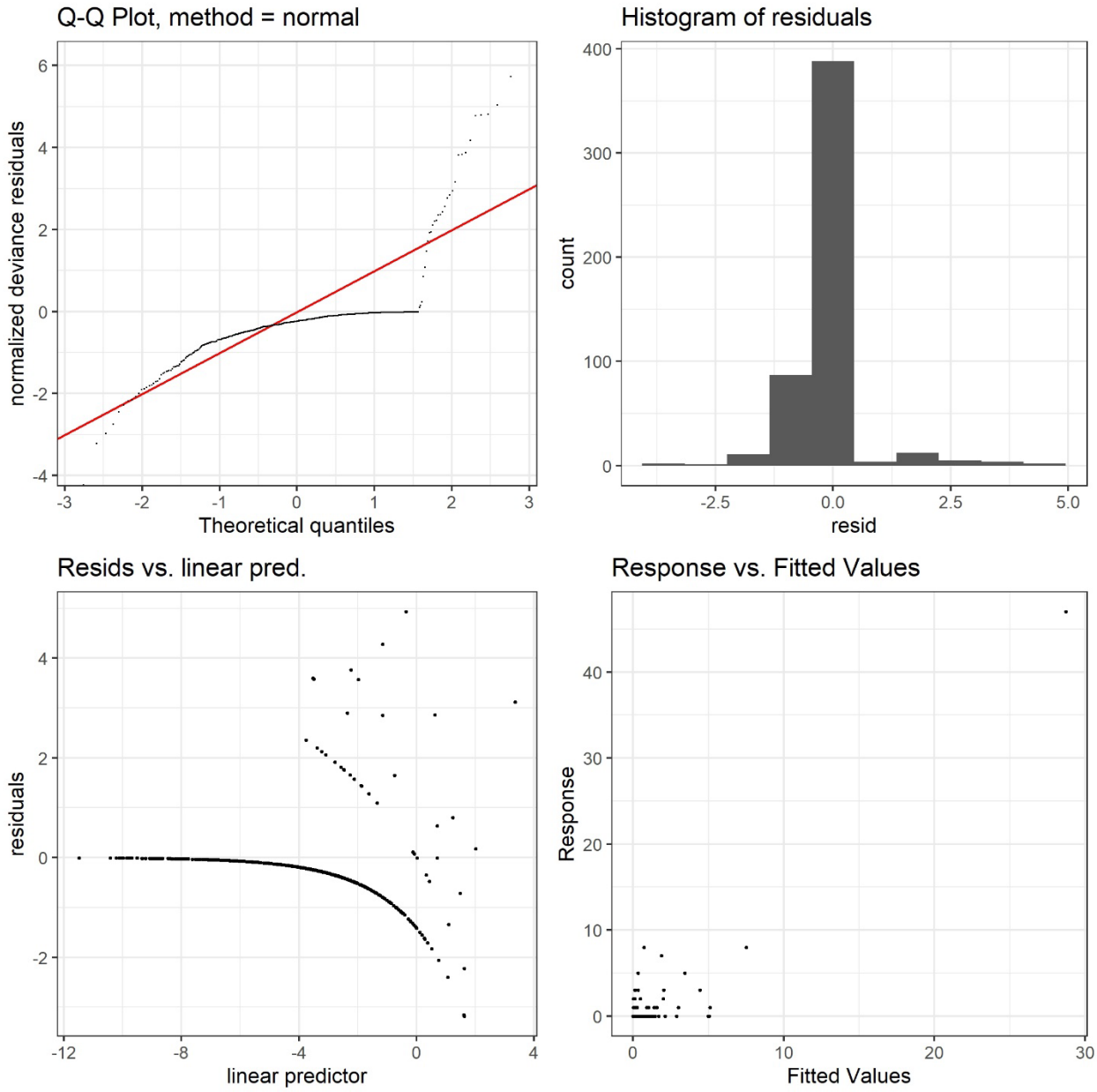


Figure A2.9. Model validation plots indicating violation of the assumptions of normality and heterogeneity of variance.

## Appendix 3. Canonical Analysis of Principal Coordinates Classification Scheme and Indicator Power Values

A3.1. Year

The CAP for year produced an overall classification success rate of 61.22% on 17 axes. We can assume that this model is significantly better than a random allocation model as the p-

value was significant (0.0047) and the measured accuracy falls within the 95% confidence range. However, classification success varied between years, with 70.92% of 2009 samples being assigned to the correct year, but only 35% of the 2011 samples being assigned to the correct year. More 2009 samples were assigned to 2012 than 2011, indicating that beta-diversity in 2009 was more similar to 2012 than 2011. More 2011 samples were assigned to 2012 than 2009, indicating that the beta-diversity in 2011 was more similar to 2012 than 2009. Finally, more 2012 samples were assigned to 2009 than 2011, indicating that the beta-diversity in 2012 was more similar to 2009 than 2011. These results are supported by the PCOA ordination (Figure 12).

Table A3.1. Taxa indicator power values and significance for year.

| *Taxon* | *Indicator Power Value* | *Group* | *p-value* |
|---|---|---|---|
| *Loliginidae* | 0.3469 | 2011 | 0.0001 |
| *Octopodidae* | 0.4600 | 2011 | 0.0001 |
| *Octopoteuthidae* | 0.0736 | 2012 | 0.0163 |
| *Pyroteuthidae* | 0.3860 | 2009 | 0.0001 |

A3.2. Season

The CAP for season produced an overall classification success rate of 70.64% on 20 axes. We can assume that this model is significantly better than a random allocation model as the p-value was significant (0.0012) and the measured accuracy falls within the 95% confidence range. However, classification success varied between seasons, with 87.22% of winter samples being assigned to the correct season, but only 8% of the spring samples being assigned to the correct season. More winter samples were assigned to early fall than spring, indicating that beta-diversity in winter was more similar to early fall than spring. More spring samples were assigned to winter than early fall, indicating that the beta-diversity in spring was more similar to winter than early fall. Finally, more early fall samples were assigned to winter than spring, indicating that the beta-diversity in early fall was more similar to winter than spring.

Table A3.2. Taxa indicator power values and significance for season.

| *Taxon* | *Indicator Power Value* | *Group* | *p-value* |
|---|---|---|---|
| *Ancistrocheiridae* | 0.0930 | Spring | 0.0091 |
| *Enoploteuthidae* | 0.2600 | Spring | 0.0284 |
| *Loliginidae* | 0.2318 | Early Fall | 0.0011 |
| *Octopodidae* | 0.3370 | Early Fall | 0.0015 |
| *Ommastrephidae* | 0.2768 | Spring | 0.0184 |
| *Pyroteuthidae* | 0.3695 | Winter | 0.0001 |

A3.3. Diel Period

The CAP for diel period produced an overall classification success rate of 59.56% on 16 axes. We can assume that this model is significantly better than a random allocation model as the p-value was significant (0.0002) and the measured accuracy falls within the 95% confidence range. Classification success was approximately equal between periods, with 58.89% of daytime samples being assigned to the correct period and 60.22% of nighttime samples assigned to the correct period.

Table A3.3. Taxa indicator power values and significance for diel period.

| *Taxon* | *Indicator Power Value* | *Group* | *p-value* |
|---|---|---|---|
| *Octopodidae* | 0.2577 | Day | 0.0466 |

A3.4. Region

The CAP for region produced an overall classification success rate of 50.42% on three axes. We can assume that this model is significantly better than a random allocation model as the p-value was significant (3.221e-07) and the measured accuracy falls within the 95% confidence range. However, classification success varied widely between regions, with 78.05% of GOM slope samples being assigned to the correct region, but 0% of the NW GOM shelf, NC GOM shelf, and GOM basin samples being assigned to the correct region. Both NW GOM shelf and NC GOM shelf samples were identified as WSF or GOM slope samples, while GOM basin

samples were identified as GOM slope or WFS samples. GOM slope samples that were not identified correctly were identified as WFS while WFS samples that were not identified correctly were identified as GOM slope.

Table A3.4. Taxa indicator power values and significance for region.

| *Taxon* | *Indicator Power Value* | *Group* | *p-value* |
|---|---|---|---|
| *Ancistrocheiridae* | 0.1150 | GOM Basin | 0.0097 |
| *Loliginidae* | 0.2194 | NC GOM Shelf | 0.0021 |
| *Ommastrephidae* | 0.3063 | GOM Basin | 0.0004 |
| *Pyroteuthidae* | 0.1934 | GOM Slope | 0.0205 |

A3.5. Depth Bin

The CAP for depth bin produced an overall classification success rate of 30.03% on three axes. We fail to reject the null hypothesis that there is no significant difference between the CAP-based model's classification success rate and that of a random allocation model (p-value = 0.2162). Classification success varied widely between depth bins, with 80.21% of 20-40 m depth bin samples being assigned to the correct depth bin, but 0% of the 0-20 m and 80-100 m depth bin samples being assigned to the correct depth bin.

Table A3.5. Taxa indicator power values and significance for depth bin.

| *Taxon* | *Indicator Power Value* | *Group* | *p-value* |
|---|---|---|---|
| *Chiroteuthidae* | 0.0491 | 100-130 m | 0.0359 |
| *Cranchiidae* | 0.1041 | 80-100 m | 0.0038 |
| *Joubiniteuthidae* | 0.0647 | 100-130 m | 0.0111 |
| *Pyroteuthidae* | 0.1409 | 100-130 m | 0.0143 |