20.6.23

2023-06-20

```
## Warning in anova.negbin(m10b): tests made without re-estimating 'theta'
```

Abstract

Sponges play a crucial role in marine ecosystems, contributing to nutrient cycling, water filtration, and habitat creation. Recent studies have also shown sponges may be a key player in the marine silica cycle. Therefore, understanding their population dynamics is crucial for unravelling ecological processes and future biodiversity studies. In this report, statistical analysis was done to analyse the population dynamics of siliceous sponge species in the Bay of Brest, France. This analysis revealed intriguing patterns and significant findings that elucidate the factors influencing sponge spatial distribution and growth. Firstly, substantial variation in siliceous sponge abundance and biomass was observed over six benthic habitats. The contrasting environmental conditions and availability of resources among habitats likely play a pivotal role in driving these variations. Additionally, the influence of predation on the abundance and biomass of the siliceous sponge *Hymeniacidonpelvis* was investigated by comparing biomass between months corresponding to predation and recovery periods. This analysis revealed a significant effect of the month on biomass aligning with before predation, after predation and recovery periods. This finding suggests a close interplay between prey-predator dynamics and sponge populations. In addition, we investigated the potential correlations between the biomass of H. perlevis and key environmental factors, including pH, temperature, and salinity. Surprisingly, this analysis indicated no significant correlation, suggesting that other factors may govern the distribution and abundance of the predator. Overall, the comprehensive statistical analysis done in this report provides valuable insights into the population dynamics of siliceous sponges. These findings contribute to a deeper understanding of marine ecosystems and lay the groundwork for further ecological studies.

1.Introduction

Sponges play a crucial role in marine ecosystems, contributing to nutrient cycling, water filtration, and habitat creation. Despite this importance, it is unclear the influencing factors of distribution, abundance, and growth. It is, therefore, crucial to understand their population dynamics including annual variation, the effect of abiotic (temperature, pH, salinity) and biotic factors (predation, etc) and habitat suitability. This report aims to enhance the knowledge of *Hymeniacidon perlevis* (*H.perlevis*) and other siliceous sponges in the Bay of Brest, France for future biodiversity assessments, conservation planning and ecological surveys.

In this report, five hypotheses were tested to gain a deeper understanding of the population dynamics of *H.perlevis*. Firstly, the differences in abundance and biomass were analysed over six benthic habitats in the Bay. By analysing this metric, more information can be deduced about the suitability of the habitats for siliceous sponges. Furthermore, predation is a biotic factor that exerts pressure on sponge populations, leading to changes in biomass. Therefore predator-prey relationships should also be considered when analysing changes in sponge biomass over periods. In the maerl beds of the bay, a spongivorous nudibranch *Doris verrucosa* (*D.verrucosa*) is present from mid-June to mid-November (1). The authors of the metadata used in this report studied the biomass of *H.perlevis* for four months, for the following reasons: June 2020 represents the population before predation, November 2020 after predation and April and June 2021 represent the sponge population recovery. It is expected that there will be a significant difference between the before and after predation periods, as well as the recovery periods.

H.pelervis biomass was also examined concerning the abiotic factors of temperature, pH, and salinity. Previous empirical studies on the species showed that the species experience a decline in biomass during colder months (2). Additionally, a substantial decrease in the pH of seawater indicates acidification, which

disrupts the physiological processes of sponges and leads to reduced growth rates and a subsequent decrease in biomass (3). Salinity is also an important parameter for studying marine species. Studies on microalgae indicate that higher salinities may influence biomass. Accordingly, this parameter was also examined.

2. Methods

Two studies and their corresponding metadata have been used in this report for assessing the population dynamics of H.perlevis, metadata from ' Sponge contribution to the silicon cycle of a diatom-rich shallow bay' done by López-Acosta (2022) and 'Nudibranch predation boosts sponge silicon cycling' also by López-Acosta (2023).

All statistical analyses were performed using the software package R version 4.2.3 (R Development Core Team 2023) (6). The packages used in R were as follows lme4, car, lme4, multicomp, ggplot2, knitr, kableExtra, MASS, and emmeans.

2.1 Study 1: Habitat study (2022 dataset)

2.1.1 Data collection

The study by López-Acosta (2022) looked at a total of 45 siliceous sponge species living in the Bay and their contribution to the silicon budget of the Bay of Brest coastal system. H.perlevis is one of the sponges accounted for in this study, therefore, this metadata was analysed. In this study, the total abundance (N sponge m^{-2}) and sponge biomass (mL sponge m^{-2}) of siliceous sponges was analysed per random quadrant (1 x 1 m), in six benthic habitats (López-Acosta et al., 2022).

The habitats were as follows:Rocky Intertidal ($_{\text{RI}}$), Rocky Subtidal ($_{\text{RS}}$), Shallow Mud ($_{\text{SM}}$), maerl beds ($_{\text{MB}}$), heterogeneous Sediments ($_{\text{HS}}$) and circalittoral coarse sediment ($_{\text{CS}}$) (refer to figure 1). Each siliceous sponge found in a quadrant's biomass was measured according to a method by López-Acosta et at.,2022. This report assesses whether there is a difference between the siliceous sponge's biomass-corrected abundance and biomass in the six habitats in the bay.
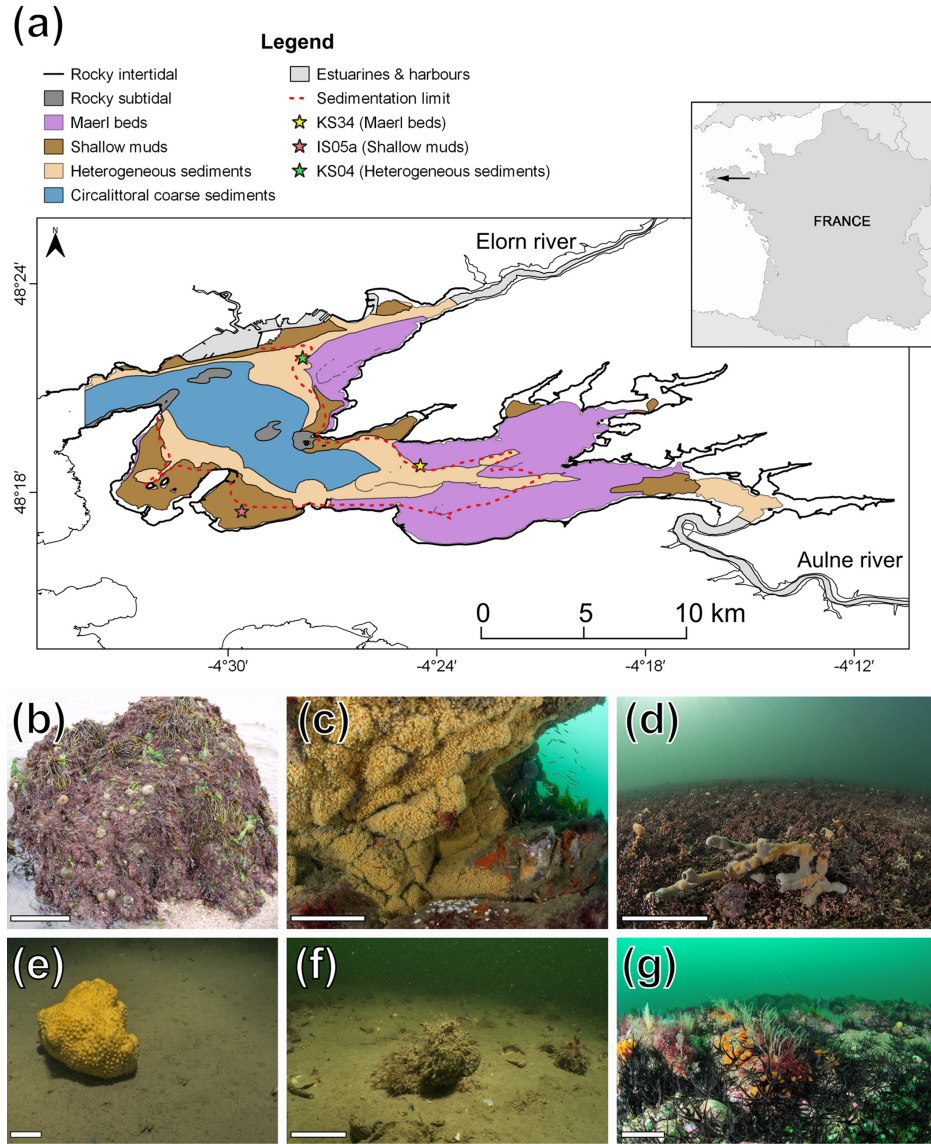
Figure 1: The six habitats in the study. (b) rocky intertidal, (c) rocky subital, (d) maerl beds, (e) shallow muds, (f) heterogenous sediments and (g) circalittoral coarse sediments.

2.1.2 Statistical analysis

2.1.2.1 The Biomass-corrected abundance of sponges per habitat (Hypothesis 1)

Independent Variable (IV): Habitat (RI, RS, SM, M, HS, CS) and biomass (mL sponge m^{-2}) Dependent (D): Abundance (N sponge m^{-2})

An Analysis of Covariance (ANCOVA) on a generalised linear model (GLM) was employed to examine the relationship between the biomass-corrected abundance of the sponges and the habitat. A homogeneity of slopes test on the model determined if an interaction between habitat and logarithmic biomass significantly affected the response variable, abundance. In cases where the interaction was not significant, Analysis of

variance (ANOVA) was performed on the GLM model.A posthoc Tukey's test was performed on the model where significant results were observed.

For this analysis, the following assumptions were made. It was assumed that no two observations in the dataset were related to each other. It was assumed that the residuals of the models followed a normal distribution (refer to figure 2). An examination of the scatter plot of deviance (refer to figure 3) residuals against the fitted values revealed a pattern suggestive of homoscedasticity. For the GLM model, the response variable is assumed to follow a negative binomial distribution, accounting for overdispersion in the model ($X^2 = 152.62$, rdf $= 135$, $p = 0.14$) (refer to table 1).

Normal Q–Q Plot
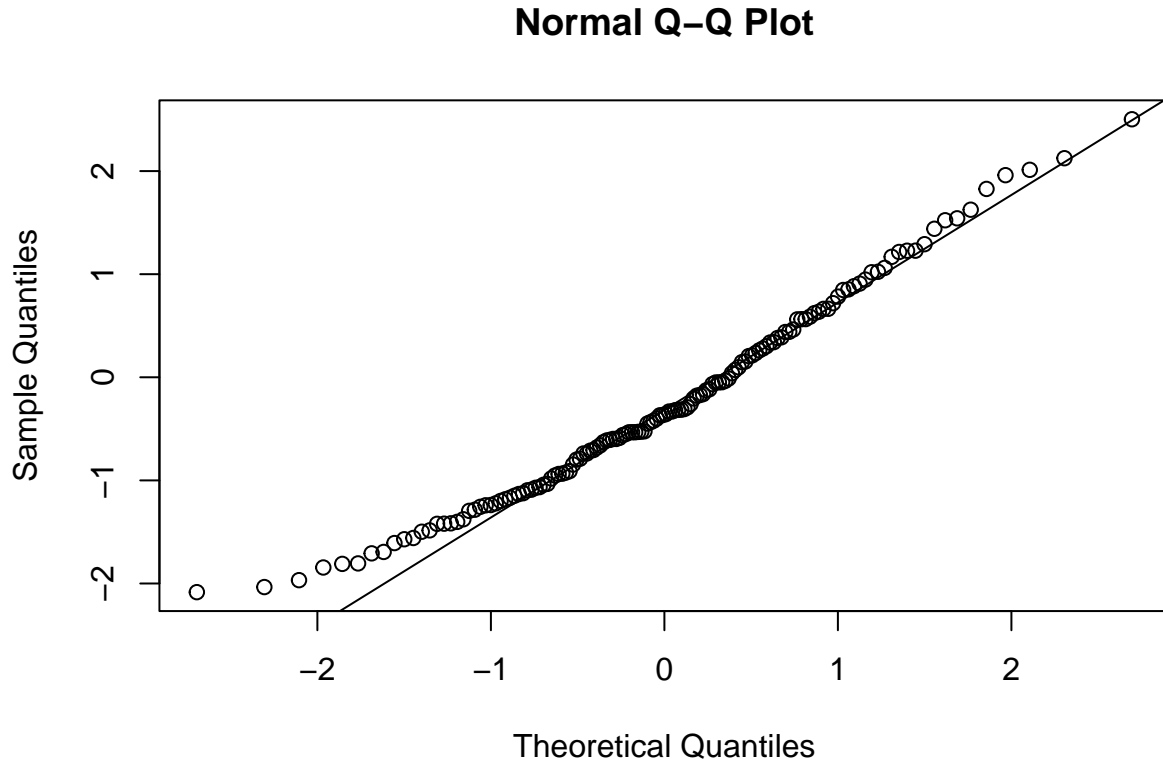
Figure 2:Q-Q plot of the distribution of the deviance residuals in the best-fitted model used for hypothesis 1

Table 1: Overdispersion of Negative Binomial Generalised linear Model (GLM), with an the predictor variables Habitat and log-biomass on the response variable abundance (hypothesis 1)

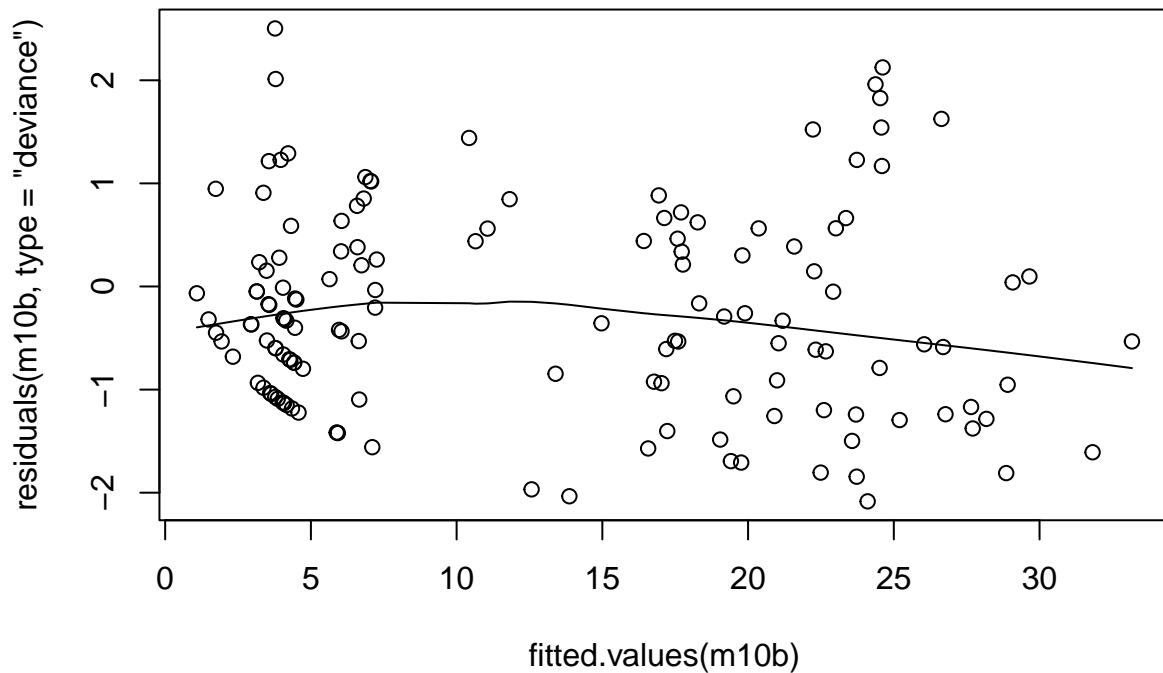|	x
chisq	152.6227650
ratio	1.1305390
rdf	135.0000000
p	0.1424801

Figure 3: Scatterplot of the predicted values versus deviance residuals plot of the best-fitted model for hypothesis 1: The negative bionmial GLM with the predictor variables habitat and log biomass on abundance (hypothesis 1)

2.1.2.2. Biomass of each sponge individual per habitat (Hypothesis 2)

IV: Habitat (RI, RS, SM, M, HS, CS) D: log-transformed biomass (mL sponge m^{-2})

An ANOVA model was used to examine the relationship between the average biomass of the sponges and habitat. Before the interpretation of the results from the ANOVA and posthoc Tukey's test the following assumptions were made. No two observations in the dataset are related to each other or affect one another in any way. The assumption of homogeneity of variance was assessed using Levene's test.This test indicated that the assumption of homogeneity of variance was met ($F_{5,136} = 0.917$, $p= 0.472$) (refer to table 2).The residuals of the data followed a normal distribution, meeting the assumption of normality (refer to figure 4).

Table 2: Levenes test on best-fitted linear model between log biomass and habitat in hypothesis 2)

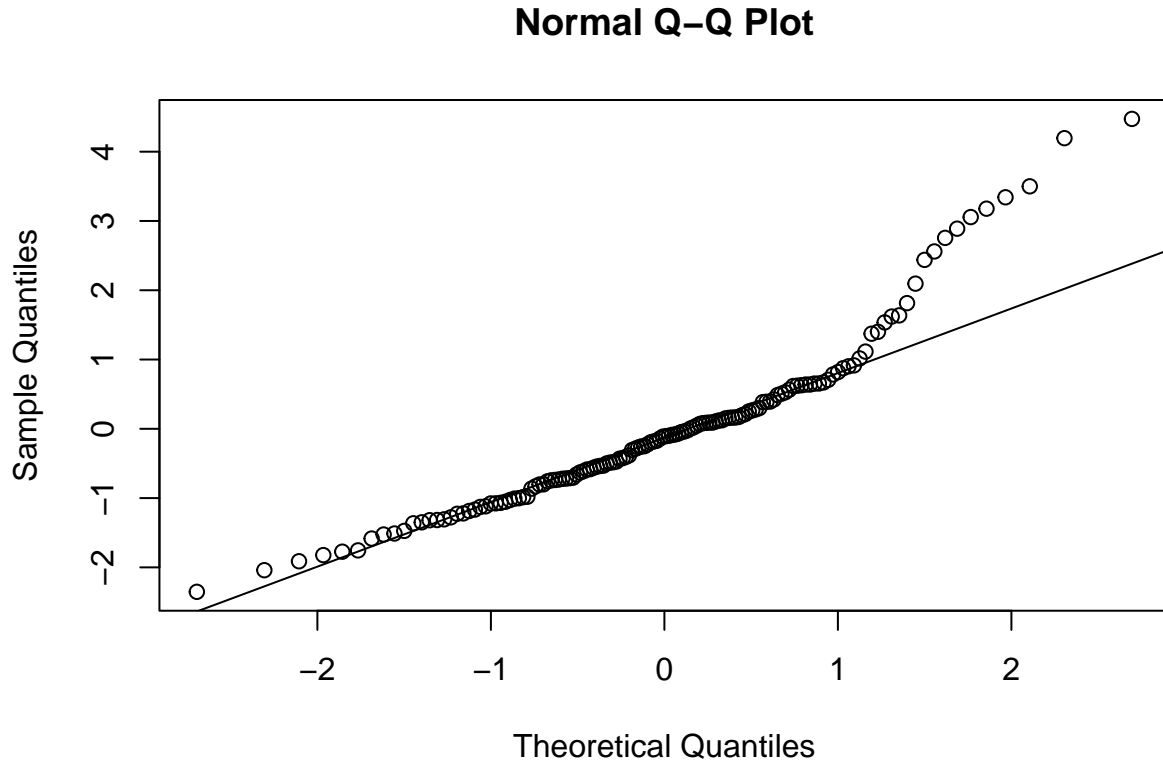|	Df	F value	Pr(>F)
group	5	0.9174826	0.4717432
	136	NA	NA

Normal Q–Q Plot

Figure 4:Q-Q plot of the distribution of the residuals in the linear model for hypothesis 2

```
## Warning in leveneTest.default(y = y, group = group, ...): group coerced to
## factor.
```

2.2 Study 2: Predation study on H.perlevis (2022)

2.2.1 Data collection

The authors (López-Acosta et al., 2023) measured the abundance and biomass of H. perlevis in randomly assigned quadrants (1 x 1 m) in the Bay over four months using the same methods as the 2022 study. Only one habitat was selected in this study, the Maerl beds. The metadata of this study was used to test the hypothesis that abundance and biomass are influenced by month according to predation periods by *D.verrucosa* During this part of the study, biomass was also considered as a factor affecting sponge abundance. The second part of this research question looked at the effect of the data on the total volume of H.perlevis per quadrant. Lastly, the effect of temperature, pH and salinity (from SOMLIT) on the biomass of H.perlevis per individual was assessed

2.2.2 Statistical analysis

2.2.2.1 The effect of date on abundance of H. perlevis (Hypothesis 3)

Table 3: Overdispersion test for best-fitted model in hypothesis 3

	x
chisq	10.4783205
ratio	0.1612049
rdf	65.0000000
p	1.0000000

Table 4: Levenes test on pearson residuals for best-fitted model in hypothesis 3

X.	Df	F.value	Pr..F.
group	3	0.8323	0.4808
	67	NA	NA

IV: Date ((June 2020 (Jun20), November 2020 (Nov20), April 2021 (Apr21), Jun 2021 (Jun21)), Log-transformed biomass (mL) *D*: Abundance (N)

The effect of the date on the biomass-corrected abundance of H.perlevis was analysed using a general linear mixed effect model (GLMM). A homogeneity of slopes test was used when appropriate to test if an interaction term was needed in the model. Tukey's posthoc test was used when significant differences in data were observed in the models.

Before the interpretation of results from the generalised linear mixed effect model (GLMM) and posthoc Tukey's test, the assumptions of these tests were assessed. No two observations in the dataset are related to each other or affect one another in any way. The dependent variable follows a Poisson distribution, with Pearson residuals normally distributed (refer to figure 5). The variances of the Pearson residuals are homogeneous as indicated by Levene's test ($F_{3,67} = 0.832$, $p = 0.480$) (refer to table 4). In this analysis, observational level random effects accounted for overdispersion ($X^2 = 10.478$, rdf= 65, ratio: 0.161, $p = 1.00$) (refer to table 3). The random effects follow a normal distribution (refer to figure 6).

Normal Q−Q Plot
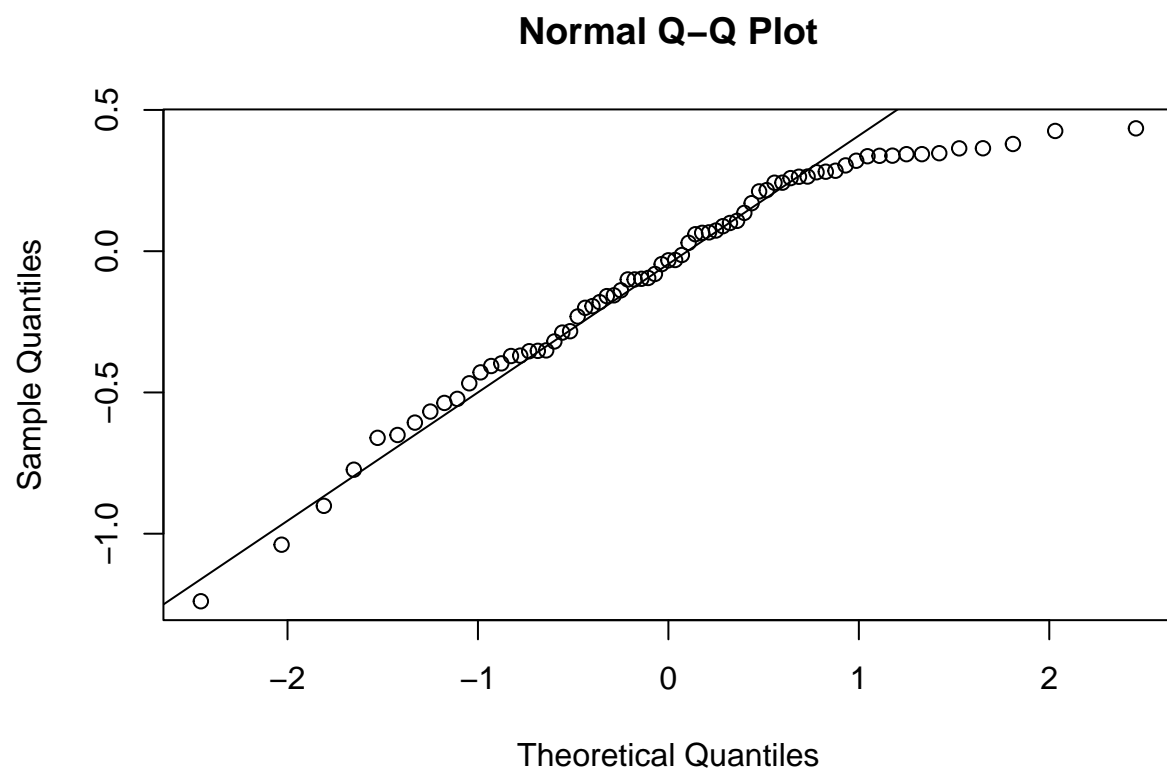

Figure 5:Q-Q plot of the distribution of pearson residuals in the linear model in hypothesis 3

Table 5: Overdispersion function for best-fitted model for hypothesis 4

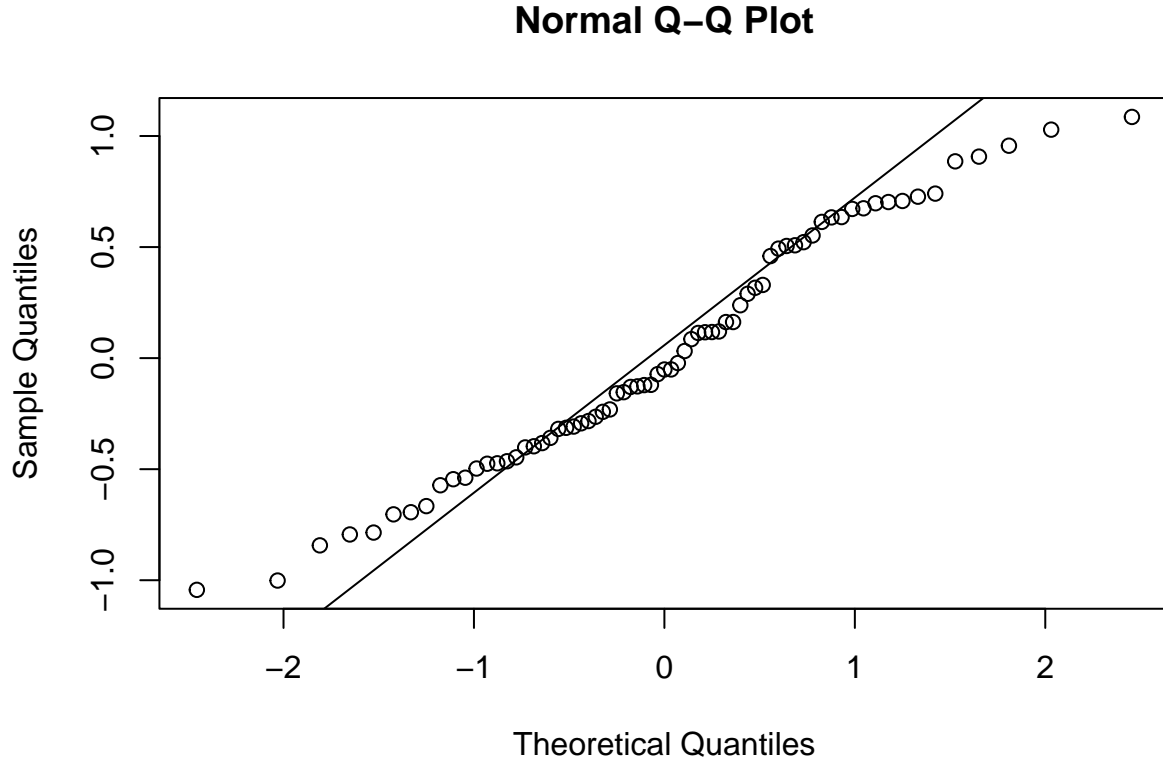|	x
chisq	35.7615256
ratio	0.5337541
rdf	67.0000000
p	0.9993809

Normal Q–Q Plot

Figure 6:Q-Q plot of the random effect in the linear model in hypothesis 3

2.2.2.2 The effect of the date on the biomass of H. perlevis (hypothesis 4)

IV: Date ((June 2020 (Jun20), November 2020 (Nov20), April 2021 (Apr21), Jun 2021 (Jun21)), *D*: Biomass (mL)

The effect of the date on the total biomass (mL) of H.Perlevis per quadrant was assessed using one-way ANOVA. When significant differences in the biomass of the sponges were observed between dates, a post hoc Tukey's Honestly Significant Difference (HSD) test was performed. The assumptions of the ANOVA and Tukey tests were as follows. No two observations in the dataset are related to each other or affect one another in any way. The assumption of homogeneity of variance was assessed using Levene's test. This test indicated that the assumption was met ($F_{3,67}$= 0.450, p= 0.718) (refer to table 6). The residuals of the tests followed a normal distribution, meeting the assumption of normality (refer to figure 7). There was no evidence of overdispersion for the final model (X^2= 35.762, rdf=67, ratio: 0.533, p= 1.00) (refer to table 5).

Table 6: Levene results on the best-fitted linear model in hypothesis 4

	Df	F value	Pr(>F)
group	3	0.4504796	0.7177922
	67	NA	NA

Normal Q–Q Plot

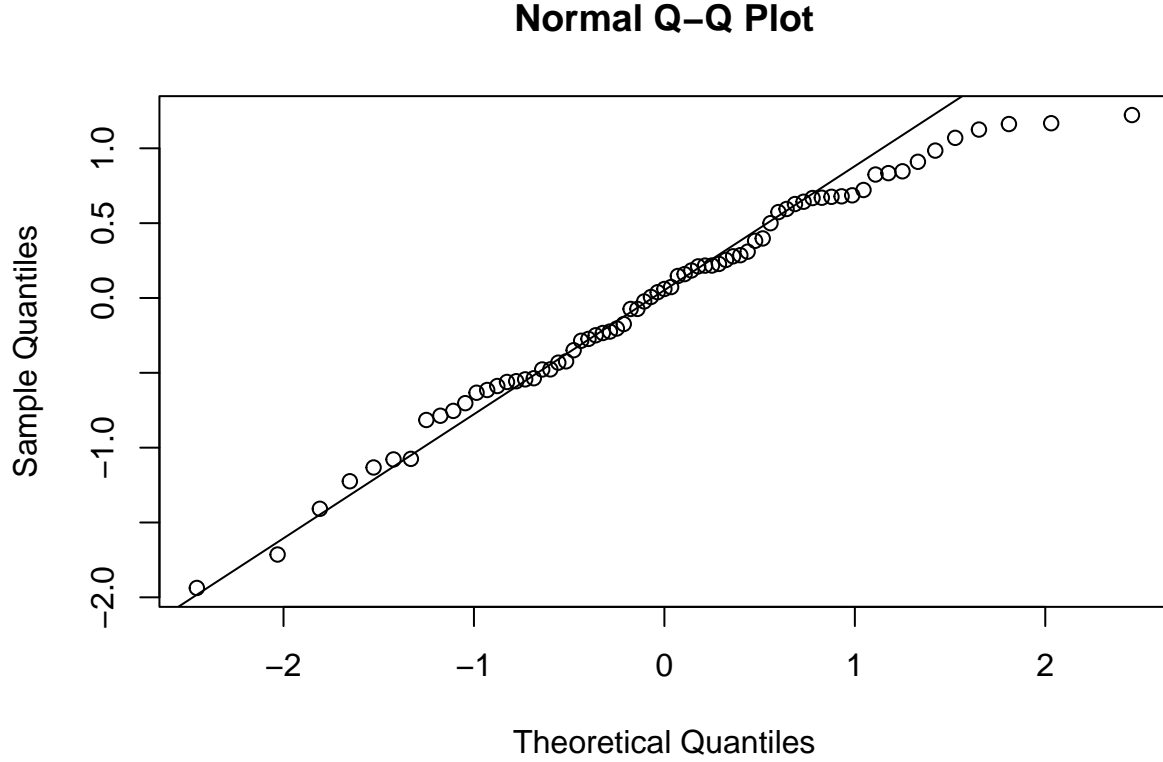*Figure 7:Q-Q plot of the residuals in the best-fitted linear model for hypothesis 4*

2.2.2.3 The effect of pH, temperature and salinity on the biomass of H.perlevis (hypothesis 5)

IV: Averaged pH, temperature (°C) and salinity (PSU) values for the dates included in this study *D*: Biomass (mL)

Pearson's product-moment correlation coefficient was employed to assess the strength and linear relationship between the average individual corrected volume (mL) of H.perlevis for the months and the same month's average pH, salinity and temperatures. The volume of each sponge was calculated by dividing the total volume per quadrant by the number of individuals. This analysis assumed that all observations in the dataset were independent, that the relationship between date and volume of H.perlevis is linear and the variance of sponge biomass is consistent across varying levels of pH, temperature and salinity.

3.Results

3.1 Biomass-corrected abundance of sponges per habitat

The highest mean ($\pm$ SD) abundance (N sponge m-2) of siliceous sponges was recorded in the Maerl beds (24.55 $\pm$ 26.17). The lowest mean abundance was recorded in the shallow mud habitat (1.60 $\pm$ 1.34) (refer to table 9).ANCOVA analysis on a GLM between biomass-corrected abundance and the habits indicated overdispersion of the model (refer to table 10). To deal with overdispersion, a negative binomial distribution

Table 7: Summary of 2022 dataset

N	Habitat	Abundance.sponges	Sponge.biomass	row_id
Min. : 1.0	Length:142	Min. : 1.00	Min. : 3.75	1 : 1
1st Qu.: 8.0	Class :character	1st Qu.: 3.00	1st Qu.: 35.21	2 : 1
Median :15.5	Mode :character	Median : 7.00	Median : 131.58	3 : 1
Mean :16.5	NA	Mean :12.77	Mean : 985.79	4 : 1
3rd Qu.:24.0	NA	3rd Qu.:15.00	3rd Qu.: 388.68	5 : 1
Max. :40.0	NA	Max. :92.00	Max. :47446.35	6 : 1
NA	NA	NA	NA	(Other):136

Table 8: Biomass mean and SD for each habitat

Habitat	Mean	SD
Circ.coarse seds	31.34526	56.54823
Heterog seds	236.67667	451.54115
Maerl	306.18825	322.05372
Rocky intertidal	123.76478	132.46652
Rocky subtidal	3779.63258	9740.55161
Shallow muds	288.60400	484.39565

was fitted for the response variable abundance. A homogeneity of slopes test was conducted to assess the interaction term between habitat and log transformed biomass (refer to table 11).This suggests that the relationship between habitat and abundance does not differ significantly across different levels of log biomass. Thus, a simpler GLM model was employed with both the predictor's biomass and habitat and their independent contribution to the prediction of abundance, this model was not overdispersed (refer to table 1).

ANOVA on the GLM denoted a significant effect of habitat ($p < 0.001$) and biomass ($p < 0.05$) on abundance (refer to table 12) Thus, a post hoc Tukey test (95% CL) was used to assess which habitats the abundance differed significantly between (refer to table 13 and 14). Significant differences were observed ($p<0.05$) between CS (p= <0.001 [-2.56,-1.15]), HS (p=<0.001 [-1.90,-0.62]), RI (p <0.001, [1.13,2.49], SM (p= <0.001 [0.99,3.98]) and MB. Significant differences were also observed between CS (p= <0.001 [-2.50,-0.87]), HS (p= <0.001 [-1.76,-0.42]), RI (p= <0.001 [-2.34,-0.93]), SM (p= <0.001 [0.83,3.79]) and RS. Figure 8 shows the variation among habitats for the abundance-biomass relationship. The largest difference between habitats for abundance was between MB and SM, corresponding to Tukey's test estimate (2.48).

```
## Warning in anova.negbin(m10b): tests made without re-estimating 'theta'
```

Table 9: Mean abundance and SD per habitat

Habitat	Mean	SD
Circ.coarse seds	4.052632	5.136727
Heterog seds	6.458333	4.934168
Maerl	24.550000	26.171966
Rocky intertidal	3.608696	2.481552
Rocky subtidal	16.387097	9.294003
Shallow muds	1.600000	1.341641

Table 10: Overdispersion test of Generalised linear Model (GLM), with an interaction term between the predictor variables Habitat and log-biomass on the response variable abundance. Overdispersion analysis of this model ($X^2 = 1415.69$, rdf $= 130$, $p = 0.00$) indicates no evidence ($p < 0.05$) to reject the null hypothesis that the model is overdispersed

	x
chisq	1415.68408
ratio	10.88988
rdf	130.00000
p	0.00000

Table 11: Homogeneity of slopes test on the interaction between preditor variables habitat and log-biomass on the response variable abundance.The likelihood ratio test (LRT) statistic was $X^2 = 10.52$ $p= 0.062$, df=5. There was no evidence ($p <0.05$) in this test to reject the null hypothesis, that the slopes of the regression lines for the log-transformed biomass and habitat are equal. In other words,there is no evidence there is an effect of the interaction between habitat and biomass in the model

	Df	Deviance	AIC	LRT	Pr(>Chi)
\<none\>	NA	140.1350	936.4396	NA	NA
Habitat:ln.biovolume	5	150.6558	936.9604	10.5208	0.0617538

Table 12: ANOVA on the best-fitted GLM model Negative Binomial

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL	NA	NA	141	269.6758	NA
Habitat	5	122.650797	136	147.0250	0.0000000
ln.biovolume	1	4.120905	135	142.9041	0.0423564

Table 13: Tukey posthoc test on ANOVA

contrast	estimate	SE	df	z.ratio	p.value
Circ.coarse seds - Heterog seds	-0.5954080	0.2892639	Inf	-2.0583560	0.3094239
Circ.coarse seds - Maerl	-1.8568538	0.2486743	Inf	-7.4670113	0.0000000
Circ.coarse seds - Rocky intertidal	-0.0481420	0.3001942	Inf	-0.1603695	0.9999854
Circ.coarse seds - Rocky subtidal	-1.6856515	0.2869418	Inf	-5.8745420	0.0000001
Circ.coarse seds - Shallow muds	0.6269637	0.5586027	Inf	1.1223785	0.8722916
Heterog seds - Maerl	-1.2614458	0.2241751	Inf	-5.6270552	0.0000003
Heterog seds - Rocky intertidal	0.5472660	0.2650648	Inf	2.0646497	0.3059961
Heterog seds - Rocky subtidal	-1.0902435	0.2345812	Inf	-4.6476171	0.0000493
Heterog seds - Shallow muds	1.2223718	0.5313871	Inf	2.3003413	0.1937604
Maerl - Rocky intertidal	1.8087118	0.2381799	Inf	7.5938891	0.0000000
Maerl - Rocky subtidal	0.1712023	0.2137222	Inf	0.8010508	0.9674069
Maerl - Shallow muds	2.4838175	0.5237780	Inf	4.7421190	0.0000312
Rocky intertidal - Rocky subtidal	-1.6375095	0.2481552	Inf	-6.5987311	0.0000000
Rocky intertidal - Shallow muds	0.6751057	0.5375282	Inf	1.2559447	0.8089073
Rocky subtidal - Shallow muds	2.3126152	0.5184902	Inf	4.4602874	0.0001193

Table 14: CI for Tukey posthoc test

contrast	estimate	SE	df	asymp.LCL	asymp.UCL
Circ.coarse seds - Heterog seds	-0.5954080	0.2892639	Inf	-1.4197248	0.2289088
Circ.coarse seds - Maerl	-1.8568538	0.2486743	Inf	-2.5655023	-1.1482053
Circ.coarse seds - Rocky intertidal	-0.0481420	0.3001942	Inf	-0.9036071	0.8073230
Circ.coarse seds - Rocky subtidal	-1.6856515	0.2869418	Inf	-2.5033510	-0.8679520
Circ.coarse seds - Shallow muds	0.6269637	0.5586027	Inf	-0.9648896	2.2188170
Heterog seds - Maerl	-1.2614458	0.2241751	Inf	-1.9002788	-0.6226127
Heterog seds - Rocky intertidal	0.5472660	0.2650648	Inf	-0.2080906	1.3026227
Heterog seds - Rocky subtidal	-1.0902435	0.2345812	Inf	-1.7587307	-0.4217562
Heterog seds - Shallow muds	1.2223718	0.5313871	Inf	-0.2919250	2.7366685
Maerl - Rocky intertidal	1.8087118	0.2381799	Inf	1.1299692	2.4874544
Maerl - Rocky subtidal	0.1712023	0.2137222	Inf	-0.4378429	0.7802475
Maerl - Shallow muds	2.4838175	0.5237780	Inf	0.9912046	3.9764305
Rocky intertidal - Rocky subtidal	-1.6375095	0.2481552	Inf	-2.3446787	-0.9303402
Rocky intertidal - Shallow muds	0.6751057	0.5375282	Inf	-0.8566913	2.2069028
Rocky subtidal - Shallow muds	2.3126152	0.5184902	Inf	0.8350709	3.7901595
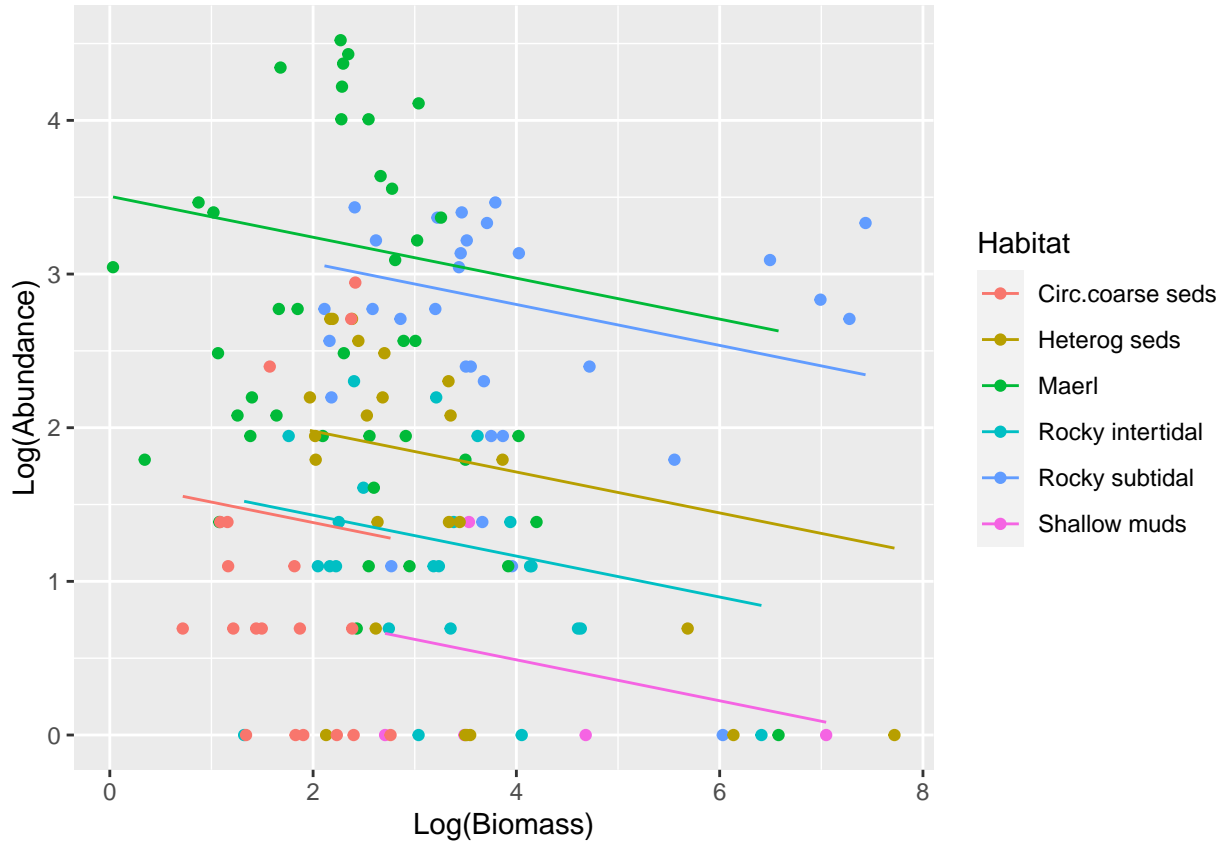
Figure 8: Plot the effect of log biomass and Habitat on the abundance of siliceous sponges in the Bay of Brest, France. Biomass is in (mL sponge m^{-2}) and Abundance is in (N sponge m^{-2})

3.2 Biomass of each sponge individual per habitat

The highest mean ($\pm$ SD) biomass (mL sponge m^{-2}) of siliceous sponges was recorded in the Rocky subtidal (3779.63 $\pm$ 132.47). The lowest mean biomass was recorded in the circalittoral coarse sediments habitat

Table 15: Mean biomass and SD per habitat

Habitat	Mean	SD
Circ.coarse seds	31.34526	56.54823
Heterog seds	236.67667	451.54115
Maerl	306.18825	322.05372
Rocky intertidal	123.76478	132.46652
Rocky subtidal	3779.63258	9740.55161
Shallow muds	288.60400	484.39565

Table 16: ANOVA on linear model with the predictor variable habitat and the response variable log biomass

Source	Df	Sum_Sq	Mean_Sq	F_value	Pr_F
Habitat	5	184.1	36.82	20.4	3.71e-15 ***
Residuals	136	245.4	1.80	NA	NA

(31.35 ± 56.55) (refer to table 15). An ANOVA model was used to examine the relationship between the logarithm average biomass of the sponges and habitat. The ANOVA model revealed a significant effect of the variable habitat on the biomass ($F_{(5,136)} = 20.4$, $p < 0.001$) (refer to table 16).To determine significant differences in biomass between habitat levels, a posthoc Tukey HSD test (95% CI) was conducted ($p < 0.05$) (refer to table 17). Significant differences were observed between the habitats HS ($p < 0.05$ [0.39,2.61]), RI ($p < 0.05$ [0.37,2,61]), RS ($p < 0.001$ [1.14,3.24]), SM ($p < 0.05$ [0.73,4.36]) and CS Significant differences were also reported between RS ($p < 0.001$ [0.69,2.42]) and SM (p = «0.05 [0.19,3.62]) and MB. Figure 9 shows the relationship between biomass and the different habitats. A large difference can be seen between CS and all other habitats besides MB, corresponding to the Tukey results.

Table 17: Tukey HSD results on ANOVA

Comparison	diff	lwr	upr	p_adj
Heterog seds-Circ.coarse seds	1.5005453	0.3896085	2.6114821	0.0020172
Maerl-Circ.coarse seds	0.6380823	-0.3699089	1.6460735	0.4501734
Rocky intertidal-Circ.coarse seds	1.4872614	0.3657042	2.6088187	0.0026023
Rocky subtidal-Circ.coarse seds	2.1890594	1.1350005	3.2431183	0.0000002
Shallow muds-Circ.coarse seds	2.5456745	0.7273083	4.3640407	0.0011977
Maerl-Heterog seds	-0.8624630	-1.7965598	0.0716338	0.0883057
Rocky intertidal-Heterog seds	-0.0132839	-1.0689265	1.0423588	1.0000000
Rocky subtidal-Heterog seds	0.6885141	-0.2951174	1.6721455	0.3345847
Shallow muds-Heterog seds	1.0451292	-0.7333382	2.8235966	0.5349857
Rocky intertidal-Maerl	0.8491791	-0.0975242	1.7958824	0.1060177
Rocky subtidal-Maerl	1.5509771	0.6853001	2.4166540	0.0000115
Shallow muds-Maerl	1.9075922	0.1915467	3.6236377	0.0199472
Rocky subtidal-Rocky intertidal	0.7017980	-0.2938130	1.6974089	0.3267299
Shallow muds-Rocky intertidal	1.0584131	-0.7267078	2.8435339	0.5250267
Shallow muds-Rocky subtidal	0.3566151	-1.3868888	2.1001190	0.9914730
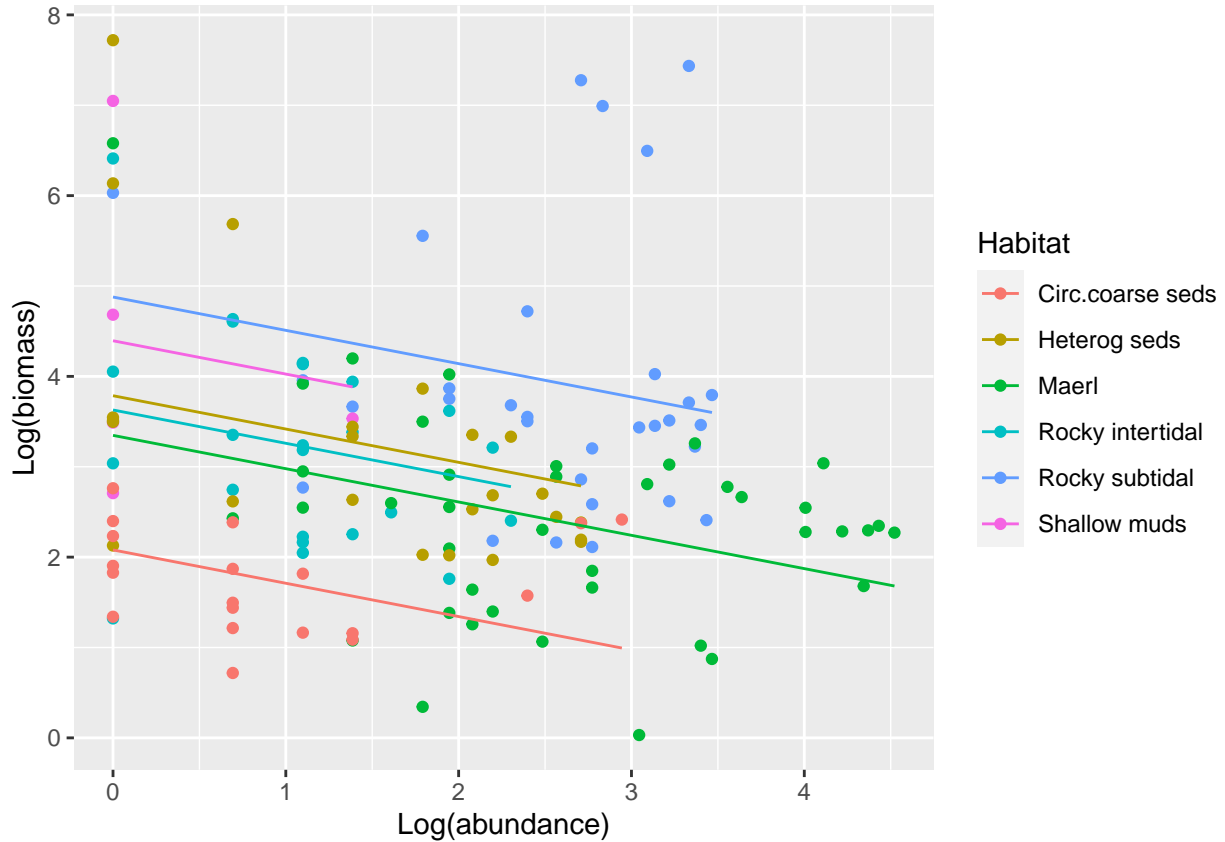
Figure 9: Plot the effect of abundance and Habitat on the log biomass of siliceous sponges in the Bay of Brest, France. Biomass is in (mL sponge m^{-2}) and Abundance is in (N sponge m^{-2})

3.3 The effect of date on abundance of H. perlevis

The mean abundance of H.perlevis over the months was 23.58 ± 16.83 individuals (N) (refer to table 18 and 19). The date with the highest mean abundance of H.perlevis (30.10 ± 23.17 individuals) was June 2021.

The date with the lowest mean abundance was April 2021 (16.87 ± 11.51 individuals) (refer to table 20).

ANOVA on a Generalised linear mixed model (GLMM) (refer to table 23 and 24) with a random grouping intercept was used to examine the relationship between size-corrected abundance and volume.Previous GLM models without the random grouping intercept were overdispersed (refer to table 22).A homogeneity of slopes test indicated that the slopes of the relationship between log-biomass and date were similar for all levels of abundance (X^2= 1.0911 p= 0.7792, df=1) (refer to table 25). Assuming a common slope of -0.4722 (refer to table 26) for all dates, the intercept differed significantly for the predictors data and volume. Tukeys posthoc test (95% CI) showed the abundance differed significantly between the dates: June 2020 – April 2021 (Pr= <0.00146 [0.34,1.85]), June2021-April2021 (Pr= 0.02585 [0.052,1.15]), November2020-June2020 (Pr= 0.0039 [2.68, -0.38]) and November2020-June2021 (Pr= 0.0033 [1.82, -0.27]) (refer to table 28 and 29). The relationship between the dates and abundance has been visualised in figure 10.
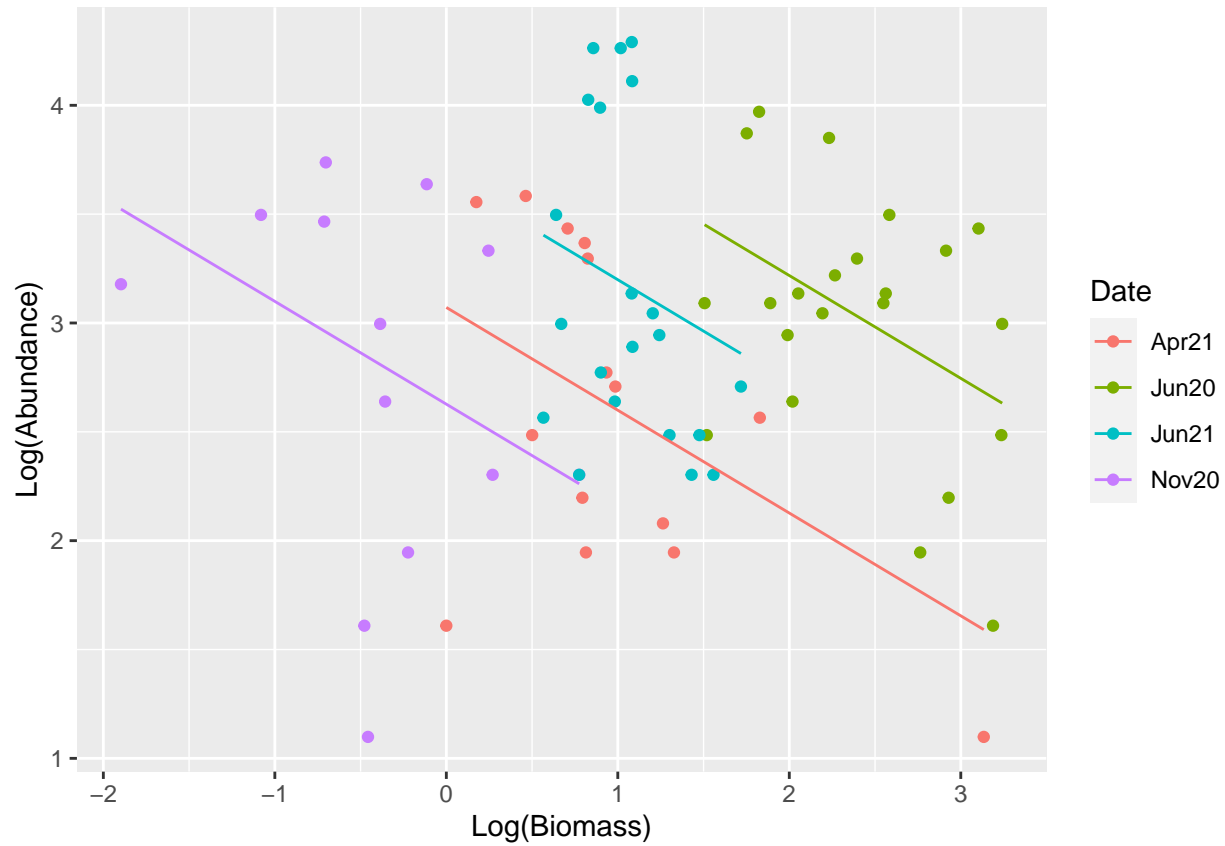

Figure 10: Plot of relationship between abundance of H.perlevis per quadrant (n) and the dates

3.4 The effect of the date on the biomass of H. perlevis

The Mean volume (± SD) of H.perlevis over the months in the study done by (López-Acosta et al., 2023) was 120.08 ± 141.53 mL (refer to table 18 a and 19) The highest sponge biomass was recorded in June 2020 (269.65 ± 164.28 mL). The lowest sponge biomass was recorded in November 2020 (14.6 ± 10.97 mL) (refer to table 21). ANOVA results indicated a significant effect (p value <0.05) of the date of biomass of the sponges ($F_{67,3}$= 51.381, p= <0.001) (refer to table 31). Post Hoc comparisons (95% confidence level) (refer to table 31) revealed that the difference between all dates was significant: Apr21-Nov20 (p.adj= 0.00032), Jun21-Nov20 (p.adj= 0.00), Jun20-Nov20 (p.adj= 0.00), Jun21-Apr21 (p.adj= 0.044), Jun20-Apr21 (p.adj= 0.00), Jun20-Jun21 (p.adj= 0.0000057).

3.5 The effect of pH, temperature and salinity on the biomass of H.perlevis

The correlation coefficient (refer to table 36) between volume and temperature at a 95% CI was r= -0.28 (-0.98, 0.88), indicating a weak negative correlation, but this observation was not statistically significant (p=

Table 18: Summary of 2023 dataset

	Var1	Var2	Freq
1		Quadrant number	Min. : 1.000
2		Quadrant number	1st Qu.: 5.000
3		Quadrant number	Median : 9.000
4		Quadrant number	Mean : 9.789
5		Quadrant number	3rd Qu.:14.000
6		Quadrant number	Max. :22.000
8		Abundance (n)	Min. : 3.00
9		Abundance (n)	1st Qu.:12.00
10		Abundance (n)	Median :20.00
11		Abundance (n)	Mean :23.58
12		Abundance (n)	3rd Qu.:31.00
13		Abundance (n)	Max. :73.00
15		Volume (mL)	Min. : 1.90
16		Volume (mL)	1st Qu.: 27.35
17		Volume (mL)	Median : 62.60
18		Volume (mL)	Mean :120.08
19		Volume (mL)	3rd Qu.:167.85
20		Volume (mL)	Max. :690.50
22		Date	Apr21:15
23		Date	Jun20:22
24		Date	Jun21:21
25		Date	Nov20:13
29		n	Min. : 3.00
30		n	1st Qu.:12.00
31		n	Median :20.00
32		n	Mean :23.58
33		n	3rd Qu.:31.00
34		n	Max. :73.00
36		row_id	1 : 1
37		row_id	2 : 1
38		row_id	3 : 1
39		row_id	4 : 1
40		row_id	5 : 1
41		row_id	6 : 1
42		row_id	(Other):65

Table 19: Standard deviation of Abundance (n) and Volume (mL) for 2023 dataset

Metric	StandardDeviation
Abundance	16.8342
Volume	141.5285

Table 20: Mean abundance and standard deviation for dates

Date	Abundance..n....Mean..	Abundance..n....SD..
Apr21	16.86667	11.51314
Jun20	23.77273	12.77621
Jun21	30.09524	23.17521
Nov20	20.46154	13.28292

Table 21: Mean volume and standard deviation

Date	Volume..mL....mean..	Volume..mL....SD..
Apr21	42.28667	22.81816
Jun20	269.65000	164.28984
Jun21	84.24762	60.01188
Nov20	14.60000	10.97490

Table 22: Overdispersion of Generalised linear Model 1: with an interaction term between the predictor variables date and log-biomass on the response variable abundance. Overdispersion analysis of this model ($X^2 = 616.73$, rdf $= 63$, $p = 0.00$) indicates no evidence ($p < 0.05$) to reject the null hypothesis that the model is overdispersed

	x
chisq	616.734174
ratio	9.789431
rdf	63.000000
p	0.000000

Table 23: Summary of Model 2:GLM with the predictor variables log-biomass and date with an interaction term between them and a random grouping intercept

Fixed.effects	Estimate	Std..Error	z.value	Pr...z..
(Intercept)	3.1512	0.2868	10.988	<2e-16 ***
ln.biovolume	-0.5586	0.2564	-2.178	0.0294 *
DateJun20	0.7878	0.6705	1.175	0.2400
DateJun21	0.9128	0.5771	1.582	0.1137
DateNov20	-0.4735	0.3547	-1.335	0.1819
ln.biovolume:DateJun20	0.1806	0.3570	0.506	0.61302
ln.biovolume:DateJun21	-0.2845	0.5233	-0.544	0.5867
ln.biovolume:DateNov20	0.2049	0.3746	0.547	0.5843

Table 24: Overdispersion of Generalised linear Model 2: with an interaction term between the predictor variables date and log-biomass on the response variable abundance. This model also includes a random grouping interept. Overdispersion analysis of this model ($X^2 = 10.76$, rdf $= 62$, $p = 1.00$) indicates there is evidence ($p < 0.05$) to reject the null hypothesis that the model is overdispersed

	x
chisq	10.7612055
ratio	0.1735678
rdf	62.0000000
p	1.0000000

Table 25: Homogeneity of slopes test for Model 2: Test on the interaction between preditor variables log-biomass and date on the response variable abundance.The likelihood ratio test (LRT) statistic was $X^2 = 1.091$ $p= 0.78$, df=3. There was no evidence ($p <0.05$) in this test to reject the null hypothesis, that the slopes of the regression lines for the log-transformed biomass and date are equal. In other words,there is no evidence there is an effect of the interaction between date and biomass in the model

	npar	AIC	LRT	Pr(Chi)
<none>	NA	571.4582	NA	NA
ln.biovolume:Date	3	566.5492	1.091062	0.7792321

Table 26: Summary of best-fitted model (M3): General linear mixed effects model (GLMM) with the predictor variables date and log-biomass on abundance. This model includes a random grouping intercept

Variable	Estimate	Std..Error	z.value	Pr...z..
(Intercept)	3.0714	0.2128	14.435	< 2e-16 ***
ln.biovolume	-0.4722	0.1428	-3.307	0.000943 ***
DateJun20	1.0911	0.2997	3.640	0.000272 ***
DateJun21	0.5996	0.2171	2.762	0.005744 **
DateNov20	-0.4443	0.3093	-1.436	0.150949

Table 27: Homogeneity of slopes test on M3

	npar	AIC	LRT	Pr(Chi)
<none>	NA	566.5492	NA	NA
ln.biovolume	1	574.8655	10.31626	0.0013186
Date	3	576.5954	16.04616	0.0011095

Table 28: Tukey posthoc test on best-fitted model (M3)

Hypothesis	Estimate	Std..Error	z.value	Pr...z..
Jun20 - Apr21 == 0	1.0911	0.2997	3.640	0.00146 **
Jun21 - Apr21 == 0	0.5996	0.2171	2.762	0.02585 *
Nov20 - Apr21 == 0	-0.4443	0.3093	-1.436	0.44277
Jun21 - Jun20 == 0	-0.4915	0.2685	-1.831	0.23402
Nov20 - Jun20 == 0	-1.5354	0.4569	-3.360	0.00400 **
Nov20 - Jun21 == 0	-1.0438	0.3071	-3.399	0.00349 **

Table 29: Tukey posthoc test confidence intervals

Variable	Estimate	lwr	upr
Jun20 - Apr21	1.09108	0.33509	1.847083
Jun21 - Apr21	0.59956	0.05204	1.147070
Nov20 - Apr21	-0.44427	-1.22451	0.335980
Jun21 - Jun20	-0.49153	-1.16881	0.185760
Nov20 - Jun20	-1.53535	-2.68788	-0.382820
Nov20 - Jun21	-1.04382	-1.81853	-0.269120

Table 30: ANOVA on LM: with predictor date and response variable log-biomass per quadrant

X.	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
Date	3	82.274	27.4246	51.381	< 2.2e-16 ***
Residuals	67	35.762	0.5338	NA	NA

Table 31: Posthoc tukey test on ANOVA

diff	lwr	upr	p.adj
Apr21-Nov20	0.4615966	1.920389	0.0003224
Jun21-Nov20	1.1747283	2.533323	0.0000000
Jun20-Nov20	2.3830388	3.729776	0.0000000
Jun21-Apr21	0.0123073	1.313758	0.0442034
Jun20-Apr21	1.2208805	2.509948	0.0000000
Jun20-Jun21	0.6151423	1.789621	0.0000057

Table 32: Summary of Temperature, pH and Salinity data from SOMLIT

ID_SITE	DATE	T	S	PH
Min. :5	Length:14	Min. :10.97	Min. :33.62	Min. :7.791
1st Qu.:5	Class :character	1st Qu.:12.90	1st Qu.:34.34	1st Qu.:7.885
Median :5	Mode :character	Median :14.24	Median :34.59	Median :7.920
Mean :5	NA	Mean :14.03	Mean :34.48	Mean :7.921
3rd Qu.:5	NA	3rd Qu.:15.62	3rd Qu.:34.73	3rd Qu.:7.950
Max. :5	NA	Max. :16.82	Max. :35.01	Max. :8.138

Table 33: Average Temperature, pH and Salinity data from SOMLIT over the months.

Month	Year	T	S	PH
6	2020	16.04400	34.68967	8.016667
11	2020	13.13850	34.46150	7.859750
4	2021	11.09100	33.85867	7.913333
6	2021	15.61075	34.81675	7.915000

0.7242, df=2). The correlation coefficient between volume and pH was r= -0.53 (-0.99,0.89), indicating a moderate negative correlation, but this observation was found to not be statistically significant (p= 0.4738, df=2). The correlation coefficient between volume and salinity was r= -0.085 (=0.97, 0.95), indicating a weak negative correlation, this observation was not statistically significant (p=0.91, df=2).

4.1 Discussion

4.1.1 Habitat study on siliceous sponges

At the habitat level, there was a significant between-habitat difference in both sponge abundance and biomass (refer to both gg plots). There was a significant difference in abundance between the MB and four other habitats in the Bay of Brest. The MB had the highest average abundance of siliceous sponges in the six habitats. Similarly, a significant difference in biomass was observed between the habitats, predominately between CS and MB with the other habitats. The CS sponge's mean biomass was significantly different from all other habitats besides MB. These results indicate that siliceous sponges may favour and reproduce better in habitats such as the Mearl beds, and sponges in RS may be larger compared to other habitats such as CS.

4.1.2 The effect of the date on the abundance

Statistical analysis revealed that the abundance of H.Perlevis (n) per quadrant differed significantly throughout the annual cycle. The highest abundance recorded was recorded in before predation and the lowest corresponds to the period after predation periods.

There was no significant difference between the dates June 2020 and June 2021, accounting for the sponge population before predation and after a year recovery. There is a significant difference in abundance between the before predation and during the recovery periods of June 2020 and April 2021. There is a significant difference between the after-predation and recovery periods in November 2020 and April 2021, June 2021.

Table 34: SD for the averaged termperature,pH and Salinity data

Month	Year	T	S	PH
11	2020	0.2453902	0.1186746	0.1302085
4	2021	0.1305412	0.2727734	0.0152753
6	2020	0.7162011	0.0520032	0.1051206
6	2021	0.5232045	0.1639296	0.0264575

Table 35: Average sponge biomass (mL) per month and year

Month	Year	Volume
6	2020	3.8781924
11	2020	12.6688840
4	2021	3.0448443
6	2021	0.8122513

Table 36: Pearsons product-moment correlation of abiotic factors on biomass

Metric	X95.CI.low	X95.CI.high	Correlation	p.value	DF
pH	-0.9877566	0.8798354	-0.5262361	0.4738	2
Salinity	-0.9670712	0.9540408	-0.0848092	0.9152	2
Temperature	-0.9777259	0.9324436	-0.2758412	0.7242	2

There was a significant difference between abundance in the before-predation and after-predation months (November2020-June2020). There was also a significant difference between the during-recovery and long-term recovery periods from June 2021-April 2021.

These results are consistent with the proposed hypothesis of this paper, regarding D.verrucosa predation. The only result not expected was the non-significant abundance observation between November 2020 (after predation) and April 2021 (5 months after the predation event). It was hypothesised that after five months the population would have recovered to a significant altitude.

4.1.3 The effect of date on biomass

There was a significant difference between the volume of sponges over the annual period. The highest sponge biomass was recorded in June 2020 and the lowest sponge biomass was recorded in November 2020 corresponding to before-predation and after-predation periods. There was a difference in the biomass of sponges over the annual period, corresponding to periods before predation, after predation and during the recovery stages. June 2020 was higher than June 2021 indicating that the biomass of sponges, may not recover fully a year after predation.

4.1.4 The correlation between pH, salinity, temp and biomass

The relationship between the average pH, salinity and temperature values and the average volume per sponge per month was found to be not statistically significant based on the data analysed. This study found that these abiotic factors do not appear to be associated with changes in biomass.

4.1.5 Limitations of this study

Several limitations should be acknowledged in the interpretation of this report's findings. Firstly, the sample size for the 2023 study was small, potentially limiting the statistical power to detect relationships proposed in this report. Secondly, the data collection period was restricted to only four months for the 2023 study, potentially overlooking seasonal variations that could influence the relationship between the studied factors. The authors did not list the time frame they conducted the 2022 study over. A study done by (2) found that the sponge species H.perlevis shows seasonal acclimation to changing environmental conditions in their morphological characteristics. Moreover, the quadrants sampled in both studies were randomly assigned, introducing the possibility of spatial variability in environmental conditions that may confound the relationship between all factors studied. Abiotic factors such as nutrient levels light availability,sediment composition and water quality were not considered in this study. These factors could have confounding effects on the volume and abundance of H.perlevis individuals. Literature has found that siliceous sponges rely on the dissolved silicon in the water to rebuild their skeletons. Therefore, silicon availability in the water may also be a confounding factor to investigate when assessing the relationship between biomass and the habitats and dates (5).

There may also be complex and synergistic effects between pH, salinity, and temperature, thus analysing each variable and their effect on volume may not be sufficient. The interaction between all three variables and their combined effect on volume should be considered in future studies. A study done (5) found that warming water temperatures may react with changes in salinity to synergistically influence oysters and other marine organisms' physiology. Finally, temperature, pH and salinity data were from surface water monitoring, which may not be representative of water conditions in the Maerl beds.

References

1. López-Acosta, M. et al. Sponge contribution to the silicon cycle of a diatom-rich shallow bay. Limnology and Oceanography 67, 2431-2447 (2022).

2. Broadribb, M., Bell, J. J. & Rovellini, A. Rapid acclimation in sponges: seasonal variation in the organic content of two intertidal sponge species. Journal of the Marine Biological Association of the United Kingdom 101, 983-989 (2021).

3. Osinga, R., Tramper, J. & Wijffels, R. H. Cultivation of marine sponges. Marine Biotechnology 1, 509-532 (1999).

4. López-Acosta, M., Potel, C., Gallinari, M., Pérez, F. F. & Leynaert, A. Nudibranch predation boosts sponge silicon cycling. Scientific Reports 13, 1178 (2023).

5. Jones, H. R., Johnson, K. M. & Kelly, M. W. Synergistic effects of temperature and salinity on the gene expression and physiology of Crassostrea virginica. Integrative and Comparative Biology 59, 306-319 (2019).

6. R Core Team (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.