# 1 Some methods

Image analysis methods were developed in the Python programming language to automatically measure the surface area of individual in each image.. The python libraries OpenCV (Bradski 2000) and ImageAI (Olafenwa and Olafenwa 2018) were used to first separate the portion of each image identified as oysters from the background and then individual oysters were automatically identified for size analysis. . Individuals were identified by automatically tracing the contour of each oyster and then estimating the surface area using a conversion factor to translate the number of image pixels occupied by each oyster to an areal size measurement using the ruler placed in each image. We analysed a subset of the same number of the oysters from each treatment. We present the workflow of different libraries in the supplementary.

Differences among treatments, experimental time points (weeks), and species for each of the morphological measurements were evaluated using generalized linear models and linear mixed-effects regression models in the R statistical programming language (RDCT 2020). The lmerTest package (Kuznetsova et al. 2017) was used to evaluate differences among treatments and weeks using a random intercept for the replicate jars assigned to each treatment. The standard glm function from the base R stats package was used when insufficient degrees of freedom prevented a random jar effect for specific subsets of the data. Models were developed separately within each week period to identify significant differences among treatments (one-way ANOVA) for each species and combined using both week and treatment as factor levels for the model predictors (two-way ANOVA as additive effects and with interactions). Post-hoc comparisons among factor levels were also conducted to compare contrasts for pairs of treatments using the estimate\_contrasts function from the modelbased package (Makowski et al. 2020).

Quantile regression analyses were also used to test for significant differences among treatments for larger individuals. This was done to account for potential differences in growth rates from density-dependent feeding effects in each jar that were independent of the treatments. Specifically, individuals in each jar that were located on the periphery of the cluster of oysters may have more access to food as compared to individuals in the center of each cluster. As a result, differences in growth between treatments may be more easily identified by evaluating only individuals with positive growth and only individuals at a higher percentile of growth. Quantile regression models were used to test for significant differences in size (area and weight) for individuals at the 80th percentile for each treatment and on the subset of individuals that showed only positive growth throughout the six-week period. The qt function from the quantreg package was used for all analyses (Koenker 2020). As for the linear models above, separate models were developed to test for differences among treatments within each week and for both week and treatment as predictors.

We visually investigated the survival of the oysters at the time of stopping the experiments and counting the number of organisms that died between the biweekly time exposure. Survival curves for each treatment and by species were estimated across the duration of the exposure period from week zero to week six using the survival package (Therneau 2020) in R (RDCT 2020). Survival rates for each of the four time points during the exposure period were evaluated using the Kaplan-Meier estimate for right-censored data (i.e, each individual was marked as 0 = alive, 1 = dead for the estimates). The survival curves produced the approximate percentage of individuals surviving at each time step with 95% confidence intervals. Significant differences between the survival curves of each treatment were then evaluated using the G-rho family of tests implemented in the survdiff function of the survival package (Harrington and Fleming 1982).

# 2 pH trends

Time series of pH at four NERRS sites were evaluated to characterize dominant frequencies of variation and their amplitudes. The goals of the analysis were to 1) evaluate if the experimental treatments were reasonable approximations of real-world variability, and 2) characterize predictable patterns of environmental conditions for inference on physiological response of biological organisms to pH variability. Sites were chosen as representative locations from the east and west coast of the United States (i.e., differences attributed to broad-scale oceanic patterns) and by latitude (i.e., north as cooler with more seasonal variation, south as warmer with less seasonal variation).

We used wavelet analysis to first identify dominant periods of cyclical periodicity in pH time series at the four NERRS sites. This produced a wavelet scalogram with period as a proportion in years (frequency, 1 equal to one year) on the y-axis and time on the x-axis, where the colors in the plot showed times with strong frequency components as warmer colors. Results from these plots allowed us to identify the relative strengths of the frequency components, such as seasonal periodicity as compared to daily or sub-daily periodicity.

Once the dominant periods were identified from the wavelet analysis, a decomposition of the time series into separate (and additive components) for long-term, daily, and sub-daily periodicity was done to quantify amplitudes attributed to each. This was done stepwise by 1) predicting the annual/seasonal variation by modelling pH vs annual time using sine wave regression (Helsel et al. 2020) and subtracting it from the observed time series, 2) predicting the daily amplitude using sine wave regression as an estimate of diel variation in pH related to biological productivity and subtracting it from the seasonally-detrended time series, and 3) predicting the remaining variation attributed to tidal forcing with harmonic regression. For the second and third steps, a subset of each time series from May to August 2012 was used as a representation of the summer period when biological productivity is likely highest at each site. For the third step, four tidal constituents were modelled to capture the dominant tidal harmonics acting at diurnal and semi-diurnal scales (O1 25.82 hours, K1 23.93, N2 12.66, S2 12, and M2 12.42). The oce R package was used for tidal harmonic regression (Kelly and Richards 2021)

The final results provided an estimate of the amplitudes of isolated variance components acting at each time scale (annual, daily, tidal). Minimum and maximum daily amplitudes for the May to August time period were also estimated as the ranges from the combined sum of the daily and tidal components. This provided an estimate of the scale of pH variation an organism may experience in any given day since the isolated variances components are additive. The separate components provided information on which frequency contributed the highest amplitude in pH and if the experimental conditions were reasonable approximations of the ranges occurring in the environment.

Kelley, D., Richards, C. 2021. oce: Analysis of Oceanographic Data. R package version 1.4-0. <https://CRAN.R-project.org/package=oce>

Helsel, D.R., Hirsch, R.M., Ryberg, K.R., Archfield, S.A., Gilrow, E.J. 2020. Statistical Methods in Water Resources. U.S. Geological Survey Techniques and Methods, Book 4, Chap. A3, 458 p. <https://doi.org/10.3133/tm4a3>