A Retrospective Analysis of Mussel Monitoring in the Puget Sound

Catherine Wangen[1,2,*] Mark D. Scheuerell[3,*]

[1] Quantitative Ecology and Resource Management, University of Washington, Seattle, WA
[2] School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA
[3] U.S. Geological Survey Washington Cooperative Fish and Wildlife Research Unit, School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA

[*] Correspondence: Catherine Wangen <cwangen@uw.edu>, Mark D. Scheuerell <scheuerl@uw.edu>

Introduction

The purpose of this report is to provide a retrospective analysis of data generated by previous mussel monitoring surveys coordinated under Washington Department of Fish and Wildlife's (WDFW) Toxics Biological Obsevation System (TBiOS). We determine how existing historical California mussel (*Mytilus californianus*) contaminant data can be used for in a Toxics in the Nearshore Vital Sign indicator. In addition, we assess the predictive ability of existing sampling rate to predict expected contaminant trends.

Toxics data was obtained by transplanting relatively uncontaminated mussels from a local aquaculture source to locations along the Puget Sound shoreline, covering a broad range of upland land-use types from rural to highly urban. Mussels were then recovered, and concentrations of several major contaminant classes were measured. Four mussels surveys were performed, with mussels being retrieved in 2013, 2016, 2018, and 2020. Our analysis focuses on polycyclic aromatic hydrocarbons (PAHs), polybrominated diphenyl ethers (PBDEs), and polychlorinated biphenyls(PCBs) due to their significance in both ecosystem and human health.

All materials used to prepare this report can be found in the following GitHub repository: https://github.com/cwangen/mussel_toxics/.

Methods

Data Analysis

The data used in this analysis originated in the form of an Microsoft Excel file, titled “2013-20MusselCagesPOPsPAHs_Cnty_WRIA_LIO_Coverages.xlsx.” The data included more fields than used in our analysis, and can be found in its entirety in “~mussel_toxics/data/raw/.” The data was cleaned in order to correct minor inconsistencies and errors, resulting in “~mussel_toxics/data/clean/totals_all.cvs.”

Dry weights of toxics found in each sample were used for analysis. Though samples did record the concentration of lipids, as well as wet weight, dry weight is the standard for reporting toxicology data. Summary tables were created for dry weights and categorical variables. Dry weight concentrations were also plotted on maps. Raincloud plots were for subcategories of WRIA and year. Raincloud plots consist of a box plot, an approximate probability density, and individual data points.

Samples from the original aquaculture source used as reference were removed except for the creation of maps. Samples outside of Puget Sound (latitude less than -123.5) were also removed from sampling, but remain in map figures. Figures 1, 2, and 3 display the concentrations of PAHs, PBDEs, and PCBs over time in mussel samples. They also shed light on how sampling is inconsistent both spatially and temporally.

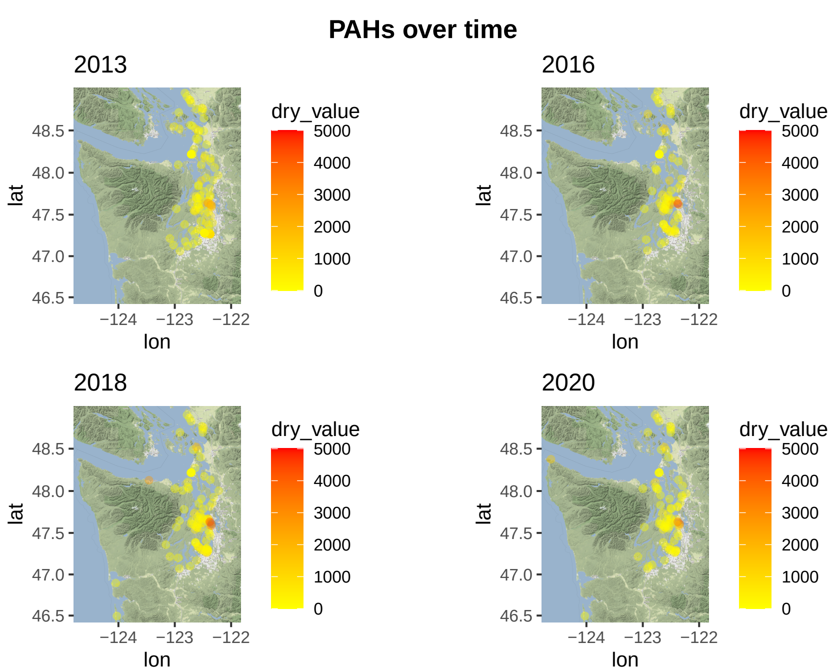

Figure 1: This one has a legend cut short due to a high value, but perhaps I should just take logs?

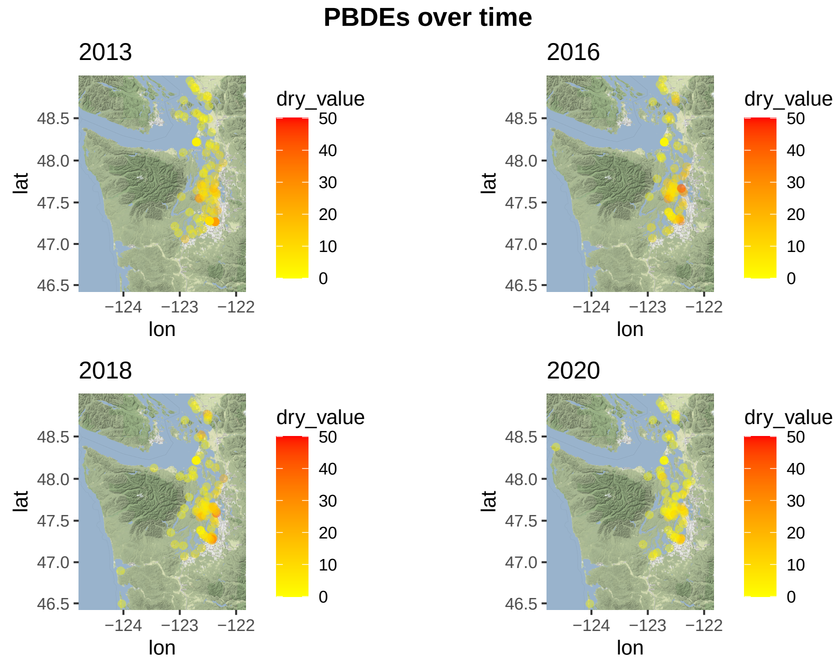

Figure 2: Logs are less of an issue here but it would be consistent.

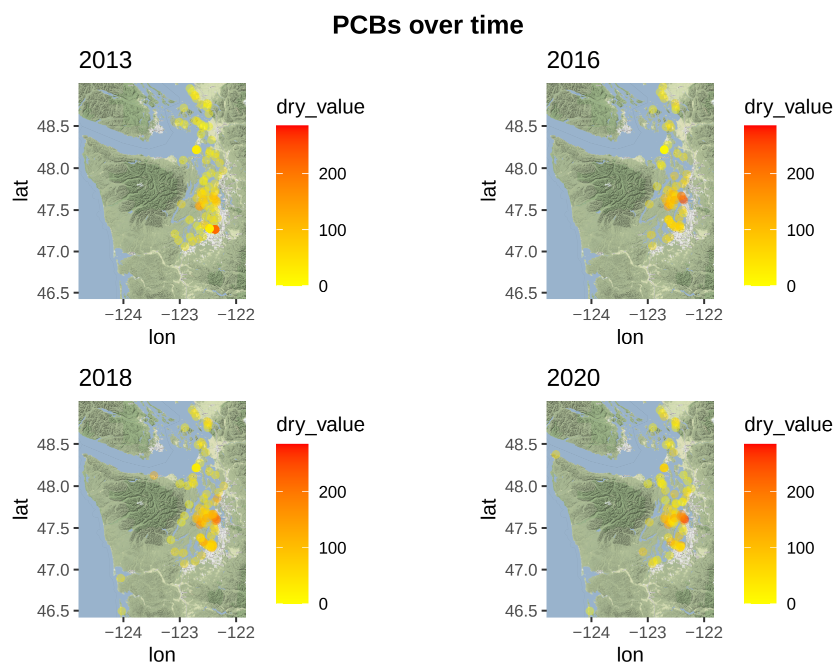

Figure 3: Again, logs? Also better titles.

Modeling

We modeled the dry weight of the analytes found in mussels using a linear mixed model (LMM). These models consist of fixed effects, which remain constant, and random effects that follow a normal distribution and can correspondingly vary by individual. Our goals were to evaluate 1) The effect of year on dry weight of the relevant toxic analyte 2) If any other factors significantly affected dry weight, and 3) the effect of WRIA by year.

Variables considered in analysis included latitude, longitude, year, county, LIO, WRIA, and mean percent AU. Other extraneous variables (IE, funding source) were not used as there is no possible relationship between these and analyte content, though this information could be helpful when we begin to focus on future sampling plans in future reports. Exploratory analysis of the data allowed us to remove extraneous factors, while including those we wished to investigate. Models were created for each of the independent analytes as their nature and distribution should not be assumed to be the same.

Initial model selection occurred by comparing Akaike information criterion (AIC) of all feasible possible models. Models that were unsuitable due to singular fits (overfitting) or violations or model assumptions such as non-normality were removed. The final model was chosen by taking into account AIC, noninfringement of model assumptions, and variables critical to project goals.

Results and Discussion

Data Analysis

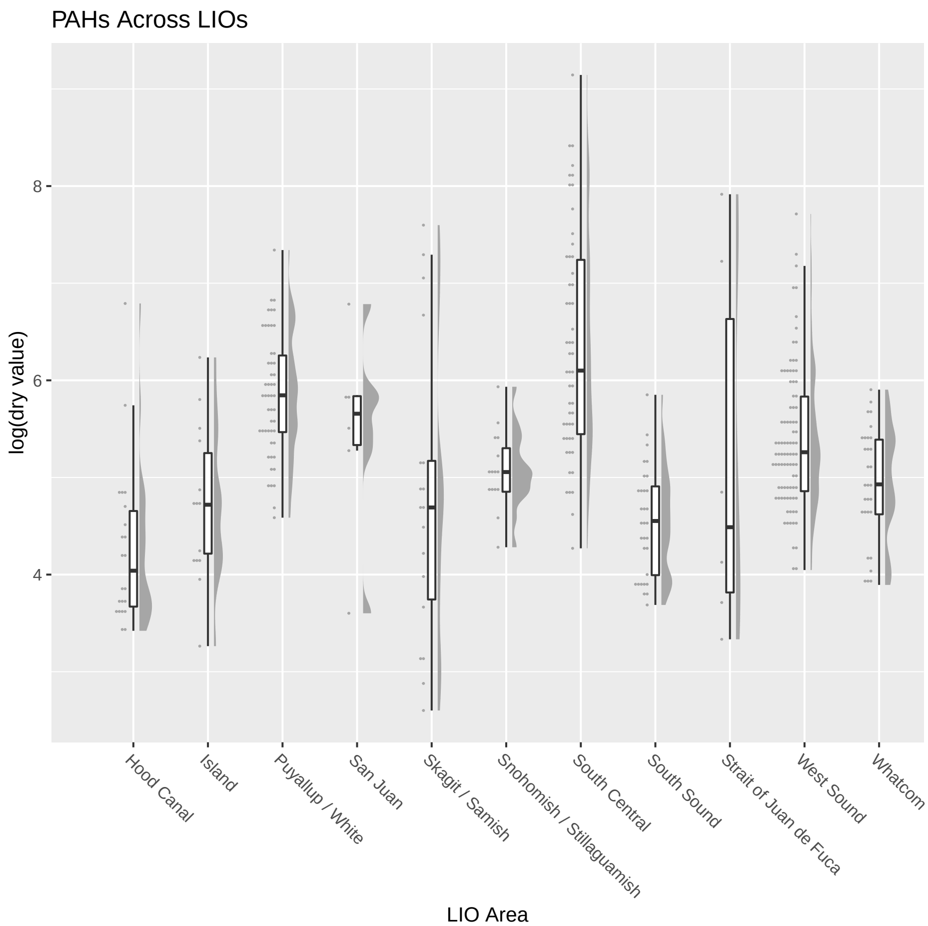

Figure 4: I figued these plots are the best to show, and the rest can go in the appendix, or just be referenced in the repo.

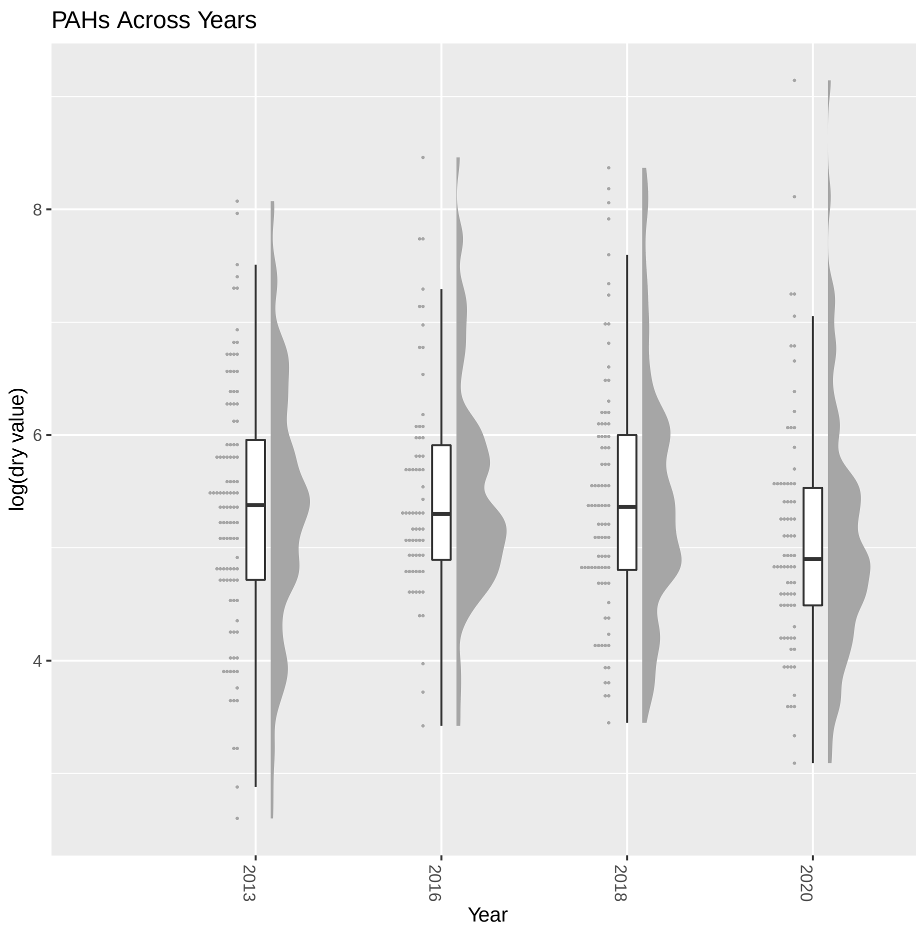

Figure 5: I figued these plots are the best to show, and the rest can go in the appendix, or just be referenced in the repo.

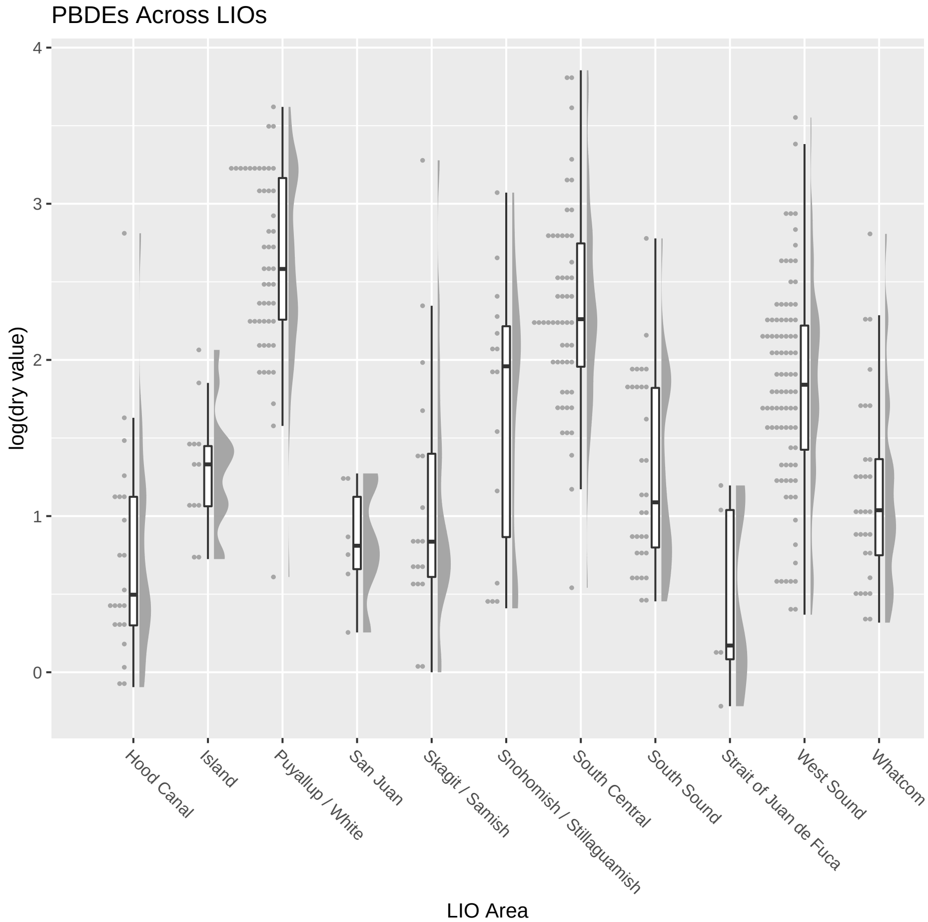

Figure 6: I figued these plots are the best to show, and the rest can go in the appendix, or just be referenced in the repo.

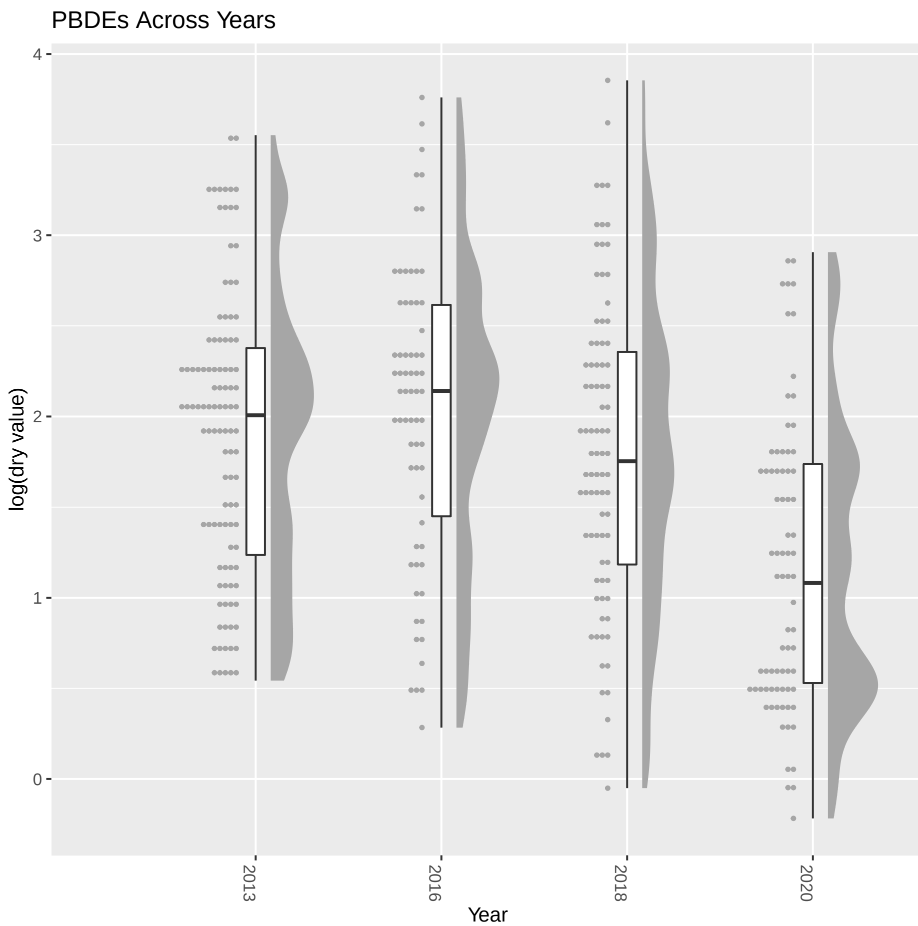

Figure 7: I figued these plots are the best to show, and the rest can go in the appendix, or just be referenced in the repo.

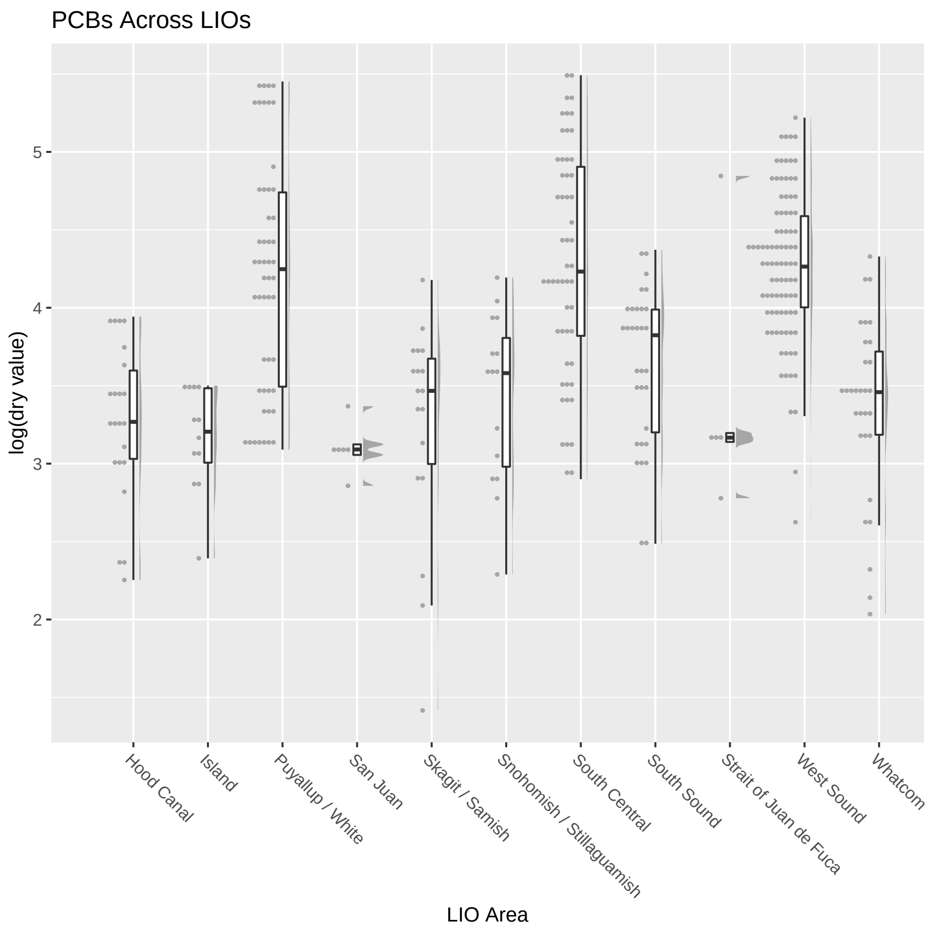

Figure 8: I figued these plots are the best to show, and the rest can go in the appendix, or just be referenced in the repo.

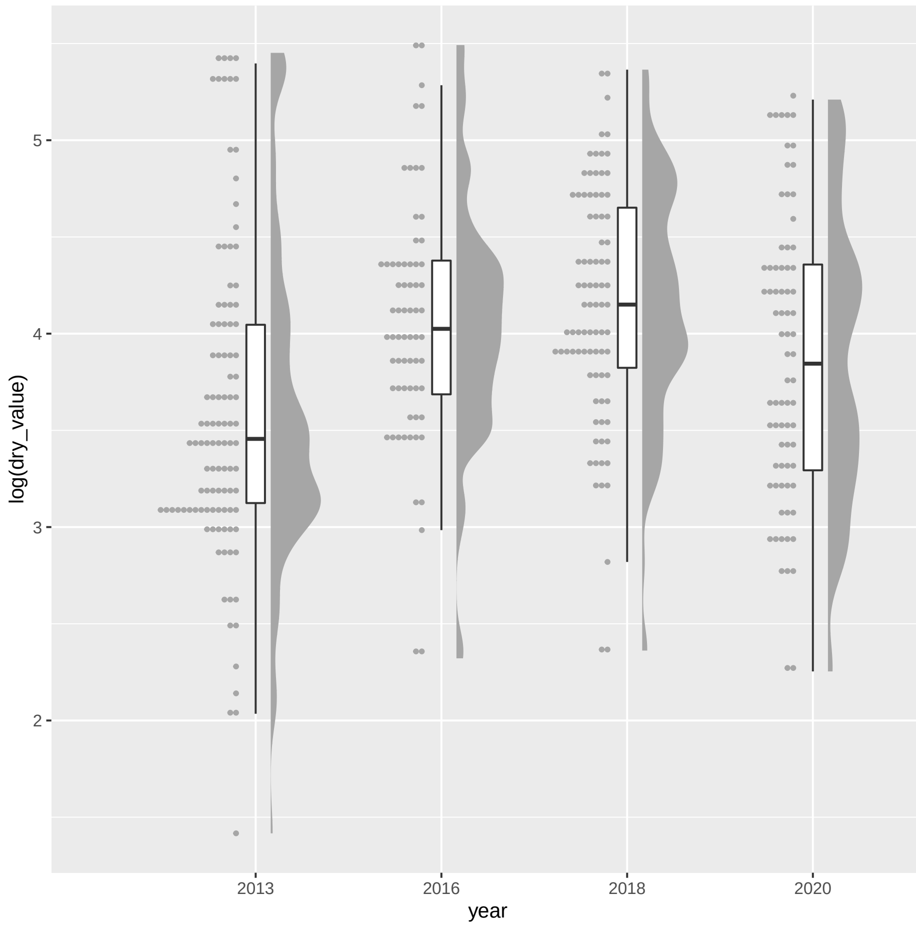

Figure 9: I figued these plots are the best to show, and the rest can go in the appendix, or just be referenced in the repo.

Modelling

The final model takes the form,

$$D_i = \beta_{1,year} + \beta_{2,year \times LIO} + \beta_3 \text{surface}_i + \beta_4 \text{time}_i + \mu_i + \epsilon_i \tag{1}$$

Where D_i is the natural logarithm of the dry weight of the analyze from a sample site, β_1 is the categorical effect of year, β_2 is the interaction effect of year and LIO, β_3 is the coefficient for mean percent AU of the nearest watershed region, β_4 is the coefficient for the time the mussels remained in the water (which varied by sampling year), μ is the random effect of longitude with $\mu_i \sim \mathrm{N}(0, \sigma^2_\mu)$, and ϵ is the error distributed $\epsilon_i \sim \mathrm{N}(0, \sigma^2_\epsilon)$.

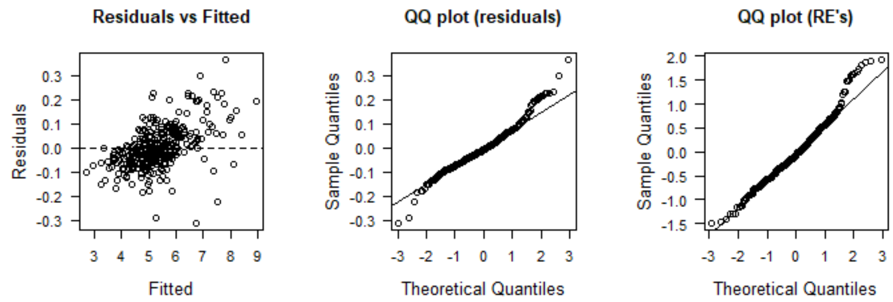

Figure 10: I figued these plots are the best to show, and the rest can go in the appendix, or just be referenced in the repo.

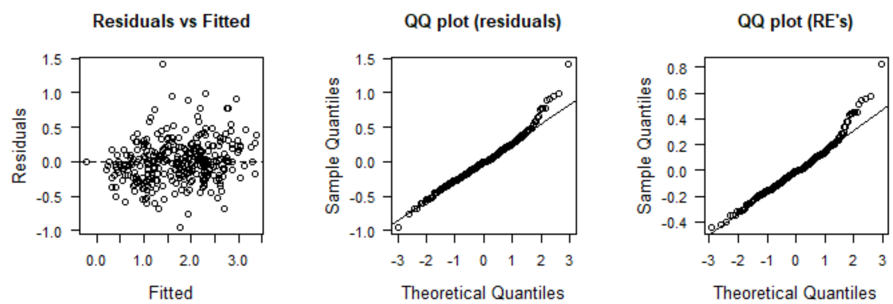

Figure 11: I figued these plots are the best to show, and the rest can go in the appendix, or just be referenced in the repo.

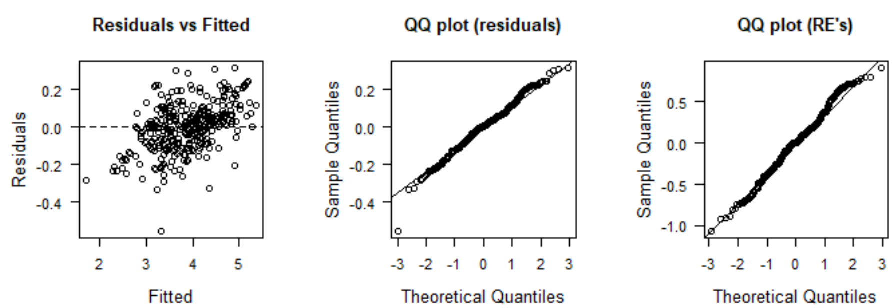

Figure 12: I figued these plots are the best to show, and the rest can go in the appendix, or just be referenced in the repo.

While we decided to fit the same model for each analyte, the parameter estimates are different in each case (REF TABLE HERE). Most notably, in each case the variable with the highest t-value is mean percent AU, indicating that it is the strongest and most significant effect on dry weight.

-model parameters (this might require a table) -partial R2 for year -estimated effect of WIRA by year (significance?), also would need a table

-what we can and can't say

Toxics in the Nearshore Recommendation

Acknowledgments

References