

## Methods

### Study Sites

I selected four sites along the California coast to measure *in situ* filtration clearance rates that represent filtration function. The sites represent a range of constructed oyster habitats including shell bag reef, shell bed, and floating aquaculture long-lines. The San Francisco Living Shorelines Project is located in northern San Francisco Bay, California (latitude, longitude: 37.964179, -122.487217; henceforth, San Rafael), and contains *Ostrea lurida* reefs constructed from July to August 2012 by The Coastal Conservancy and collaborators (Environmental Science Associates 2014, Latta & Boyer 2015). San Rafael consists of a shell bag reef matrix with a 32 m x 10 m footprint with three rows of eight shell bag reefs constructed as *O. lurida* habitat (Environmental Science Associates 2014). Each reef unit measured 2 m x 2 m x 1 m and was composed of four shell bag elements measuring 1 m x 1 m x 1 m. A shell bag was made of plastic mesh filled with clean, commercially-grown *Crassostrea gigas* shell. The matrix was positioned approximately 200 m from shore on a mudflat at -0.3 m MLLW (Mean Low Lower Water) tidal elevation (Environmental Science Associates 2014).

The Upper Newport Bay Living Shorelines Project is located in Newport Bay, California and contains restored *O. lurida* beds. It was constructed in May 2017 by Orange County Coast Keeper, and collaborators at CSU Fullerton and CSU Long Beach (Wood 2018). My research included two sites: Shellmaker (33.622097, -117.892399) and Deanza (33.620291, -117.897692) that are about 550 m apart. The oyster beds measure 20 m x 1 m x 0.25 m and were constructed with coconut coir bags filled with clean, commercially-grown *C. gigas* and *Mytilus galloprovincialis* mussel shell. The beds were positioned approximately 30 m from shore on a mudflat at approximately -0.15 m MLLW (Wood 2018).

Morro Bay Oyster Company (MBOC) is a commercial aquaculture operation in Morro Bay, California (35.334707, -120.844000; henceforth, Morro Bay). MBOC grows *C. gigas* in plastic mesh bags measuring about (80 cm x 55 cm) attached to floating lines that stretch across an area about 75 m x 75 m. The oysters grown in the bags are approximately five to eight cm in length. The bags of oysters rest on the underlying mudflat at around 0 m MLLW, and float to the surface as the tide rises.

### Experimental Setup

*In situ* filtration methods were adapted from Grizzle et al.'s (2006, 2008) upstream-downstream measurements of bivalve beds on the North American Atlantic coast. Two identical water quality sondes (Yellow Springs Instruments 6600EDS) measured the change in chlorophyll  $\alpha$  (Chl  $\alpha$ , fluorometer), temperature, salinity (automatically calculated from conductivity/temperature sensor), and turbidity (optical sensor) from positions

upstream and downstream of the oyster habitat. The sondes were hung inside of freestanding PVC (polyvinyl chloride) housings with water flow slats (Figure 1) and were set at a height to align the sensors with the approximate height of the oyster beds, reefs, or floating bags.

Prior to each filtration trial, water tracing dye (Rhodamine WT) released upstream of the habitat provided a visual indicator of water flow direction and interfering currents or eddies and was recorded by a drone about 25 feet above the water. Trials were conducted close to low tide on either the ebb or flood tide depending on how the water flow direction transected the oyster habitat. The sonde housings were positioned to measure an uninterrupted linear water flow across the oyster habitat based on the dye. Linear water flow assumes that the upstream sonde is measuring the same water as the downstream sonde. An electromagnetic meter (Marsh-McBirney Flo-Mate 2000) measured water velocity at the depth of the sensors at the beginning, middle, and end of each trial. Water velocity measurements were taken at the downstream sonde position and in the middle of the transect, velocity was averaged across positions and time. Mean depth was determined by markings on the PVC sonde housings at the beginning and end of the trial, and distance between sondes was determined post-trial with a transect tape. The sondes recorded measurements every one second, and trials lasted as long as water flow direction was consistent; ranging from six and 40 minutes.

I took the mean of temperature, salinity, turbidity of the upstream sonde during the filtration trial for analyses, and the mean of Chl  $\alpha$  at both upstream and downstream sondes for analyses. To compare the clearance rates of surrounding mudflat habitat to restored *O. lurida* and *C. gigas* aquaculture, I conducted a mudflat control trial at each site. Control trials used the same experimental set up, except the instruments were positioned over the mudflat adjacent to the oyster habitat. Data corresponding with field disturbances was cut from the time series. Field work was conducted from February 2018 to June 2019 and measured 22 oyster habitat filtration trials and four control mudflat trials.

## **Chlorophyll $\alpha$ Corrections**

Chl  $\alpha$  is a photosynthetic pigment whose concentration is a proxy for phytoplankton and macroalgal detritus consumed by filter-feeding bivalves, and is commonly used to measure filtration (Harsh & Luckenbach 1999, Grizzle et al. 2006, 2008, Wasson et al. 2015, Milbrandt et al. 2015). I compared Chl  $\alpha$  sensors readings by conducting two side-by-side trials before and after filtration and control trials. In the side-by-side trials, I placed the sondes adjacent to one another to compare the sensors' Chl  $\alpha$  concentration readings in the same mass of water for approximately 10 minutes. All data, in filtration, control, and side-by-side trials, was reviewed for extreme Chl  $\alpha$  and turbidity spikes indicating disturbances to the trial. If extreme Chl  $\alpha$  and turbidity values were corroborated by field notes, the time period was removed from the trial. Next,

I corrected mean Chl  $\alpha$  values in the filtration and control trials with the mean Chl  $\alpha$  difference from the corresponding side-by-side trials. The mean difference between sondes in combined side-by-side trials was divided in half to produce a correction term. The sonde with higher mean Chl  $\alpha$  in the side-by-side trials was corrected in the filtration and control trials by subtracting the correction term from mean Chl  $\alpha$  measurements, and the lower sonde was corrected by adding the correction term.



Figure 1: Two identical YSI 6600EDS water quality sondes were used to measure chlorophyll  $\alpha$ , temperature, salinity, and turbidity. The sondes were hung inside PVC housings to adjust the sensor depth in the water column. Water velocity measurements were taken with an Marsh-McBirney Flo-Mate 2000. Photo taken at Deanza, Newport Bay, California.

## Seston Content

I determined seston total particulate matter (TPM), particulate inorganic matter (PIM), particulate organic matter (POM), and organic content (OC) gravimetrically. I collected water samples immediately adjacent to the experimental area during the filtration or control trials, and filtered the water through a borosilicate glass microfiber filter (Whatman 9907-047 pre-wash and pre-weighed) in the field. The filters were stored on ice until they could be moved to a freezer. Thawed filters were rinsed with a 0.5 M ammonium formate solution to remove salt. Typically this step immediately precedes filtering (Gray & Langdon 2018), but this was not possible to do in the field. Filters were then dried in an oven at 60°C for 48 hours, and weighed to determine TPM (Equation 1). Next, the dried filters were ashed at 450°C for  $\geq 4$  hours and weighed to determine PIM (Equation 2) and POM (Equation 3). Seston OC is simply the ratio of POM to TPM. Gray & Langdon (2018) used washed and ashed filters; however, the filters I used that were only washed. The mean weight difference between ashed filters and washed only filters was 0.5%, which I used to corrected the filter pre-weights in Equations 1 & 2. In addition, control filters, filtered with distilled water, were stored and processed with every sampling batch. The TPM and PIM of control filters represented processing contamination within each

sample batch, and I used these values to correct sample TPM and PIM within each batch.

$$TPM(mg/L) = \frac{[FilterDryWeight(g) - FilterPreWeight(g)] \times \frac{1000mg}{L}}{Watersamplevolume(mL) \times \frac{1L}{1000mL}} \quad (1)$$

$$PIM(mg/L) = \frac{[FilterAshWeight(g) - FilterPreWeight(g)] \times \frac{1000mg}{L}}{Watersamplevolume(ml) \times \frac{1L}{1000mL}} \quad (2)$$

$$POM(mg/L) = TPM(mg/L) - PIM(mg/L) \quad (3)$$

## Filtration Calculations

I paired upstream and downstream Chl  $\alpha$  measurements in the experiment and control trials to determine how Chl  $\alpha$  concentrations changed across the habitat. I used mean water velocity and sonde distance to estimate the time it took a parcel of water to travel from the upstream sonde to the downstream sonde. The travel time was subtracted from the time-stamps of each downstream Chl  $\alpha$  measurement to pair with upstream Chl  $\alpha$  measurements; unpaired measurements were discarded. Next, I calculated percent Chl  $\alpha$  removal ( $Chl_{rmd}$ , Equation 4; Grizzle et al. 2008) and habitat clearance rate (HCR, Equation 5; Milbrandt et al. 2015) using average Chl  $\alpha$  concentrations and habitat dimensions instead of biomass.

$$Chl_{rmd} = \frac{Chl_{up} - Chl_{down}}{Chl_{up}} \times 100 \quad (4)$$

$$HCR(Lhr^{-1}m^2) = \frac{A_{Xsec} \times V \times 1000}{A_{habitat}} \times \frac{Chl_{up} - Chl_{down}}{Chl_{up}} \quad (5)$$

$A_{Xsec}$  is the cross-sectional area of the water column (mean water depth  $\times$  assumed 1 m width),  $V$  is the mean water velocity (m/hr),  $A_{habitat}$  is the area being measured (distance between the instruments  $\times$  assumed 1 m width),  $Chl_{up}$  is the mean upstream Chl  $\alpha$  concentration, and  $Chl_{down}$  is the mean downstream Chl  $\alpha$  concentration.

## Analysis

All data wrangling and analysis was conducted R 3.6.3 (R Core Team 2020, tidyverse R package; Wickham et al. 2019) and plotted using the ggplot2 package (Wickham 2016); the code is available in Github repository “Oyster-Insitu-Filtration”. Two field days (Morro Bay 2019-05-18 and Deanza 2018-10-25; Table 1) from my data set were missing TPM and OC values. I used a semi-parametric imputation to estimate the missing values

(missForest R package; Stekhoven & Bühlmann 2012). Differences in temperature, salinity, turbidity, and Chl  $\alpha$  among sites was investigated using a one-way ANOVA and post-hoc Tukey honestly significant difference (HSD) on significantly results. Percent Chl removal among sites was examined using a Kruskal-Wallis one-way analysis of variance.

To determine if there was a difference between mudflat control and oyster habitat filtration trial HCRs, I first used a two-sample, single-tail, t-test. To account for the effects of water quality (i.e. temperature, salinity, turbidity, TPM, OC) and site variables, I fit a random forest regression (including all filtration and control trials) to HCR with temperature, salinity, turbidity, TPM, OC, and site as predictor variables (randomForest R package Liaw & Wiener 2002). Control and filtration HCR were adjusted by the residuals of the random forest regression to control for the effects of water quality and site and were compared by a second two-sample, single-tail, t-test. The effect of site on filtration trial HCR was examined using a Kruskal-Wallis analysis. Another Kruskal-Wallis analysis examined the top 0.5 quantile of filtration trial data at each site to determine if the top HCR performance differed among sites. A quantile analysis ( $\tau = 0.5, 0.9$ ) was used to see if water quality variables had differential relationships with HCR.

To examine the effect of water quality variables and site on HCR, I fit a second random forest regression that only included filtration trials (randomForest R package; Liaw & Wiener 2002). The random forest model I fit is an ensemble of 2,000 individual decision trees. Each tree fits a random subset of the data and is prone to over-fitting. The final output (forest) is the average of the 2,000 individual trees, which corrects for the individual trees' weaknesses (i.e. ensemble learning, Sagi & Rokach 2018). I chose to use a random forest regression for several reasons. First, random forest regressions are non-parametric models that learn non-linearity relationships without explicitly modeling them (Grömping 2009). This works well for my data as previous research suggests that temperature, salinity, and TPM have non-linear relationships with *O. lurida* and *C. gigas* clearance rates (Gray & Langdon 2018). Second, random forest regressions work with missing data (Stekhoven & Bühlmann 2012), and I had two trials with missing TPM and OC values because of a processing error. Third, I had a large number of variables relative to observations; random forests regressions work well with this 'wide' data structure (Grömping 2009).

The relationship between TPM and OC was estimated using a linear model, as well as the relationship between *O. lurida* shell length at Deanza and DTW. Data corrections and statistical analyses were conducted in consultation with Dr. Kevin Nichols (CSUF Statistical Consulting Unit).

Environmental Science Associates (2014) San Francisco Bay Living Shorelines Project Physical Properties Report, November 2014. Environmental Science Associates.

Gray MW, Langdon CJ (2018) Ecophysiology of the Olympia Oyster, *Ostrea lurida*, and Pacific Oyster, *Crassostrea gigas*. *Estuaries and Coasts* 41:521–535.

Grizzle RE, Greene JK, Coen LD (2008) Seston removal by natural and constructed intertidal eastern oyster (*Crassostrea virginica*) reefs: A comparison with previous laboratory studies, and the value of in situ methods. *Estuaries and Coasts* 31:1208–1220.

Grizzle RE, Greene JK, Luckenbach MW, Coen LD (2006) A New In Situ Method for Measuring Seston Uptake By Suspension-Feeding Bivalve Molluscs. *Journal of Shellfish Research* 25:643–649.

Grömping U (2009) Variable importance assessment in regression: Linear regression versus random forest. *American Statistician* 63:308–319.

Harsh D, Luckenbach MW (1999) Materials Processing by Oysters in Patches: Interactive Role of Current Speed and Seston Composition. In: *Oyster reef habitat restoration: A synopsis and synthesis of approaches*. Luckenbach MW, Mann R, Wesson JA (eds) Virginia Institute of Marine Sciences, College of William & Mary, Gloucester Point, Virginia, pp 251–265

Latta M, Boyer K (2015) San Francisco Bay Living Shorelines Project: Nearshore Linkages Project Summary of Key Findings Two Years Post-installation.

Liaw A, Wiener M (2002) Classification and Regression by randomForest. *R News* 2:18–22.

Milbrandt EC, Thompson M, Coen LD, Grizzle RE, Ward K (2015) A multiple habitat restoration strategy in a semi-enclosed Florida embayment, combining hydrologic restoration, mangrove propagule plantings and oyster substrate additions. *Ecological Engineering* 83:394–404.

R Core Team (2020) R: A language and environment for statistical computing.

Sagi O, Rokach L (2018) Ensemble learning: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 8:1–18.

Stekhoven DJ, Bühlmann P (2012) Missforest-Non-parametric missing value imputation for mixed-type data. *Bioinformatics* 28:112–118.

Wasson K, Zabin C, Bible J, Briley S, Ceballos E, Chang A, Cheng B, Deck A, Grosholz T, Helms A, Latta M, Yednock B, Zacherl D, Ferner M (2015) A Guide to Olympia Oyster Restoration and Conservation: Environmental Conditions and sites that Support Sustainable Populations.

Wickham H (2016) *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag, New York.

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H (2019) Welcome to the tidyverse. *Journal of Open Source Software* 4:1686.

Wood V (2018) THE EFFECT OF SEDIMENTATION ON OYSTERS (*OSTREA LURIDA*) ADJACENT TO EELGRASS BEDS (*ZOSTERA MARINA*). PhD thesis, California State University Fullerton