

# Shelton et al. 2021 Supplement S1

## Environmental DNA analysis

This section provides detailed information about the analysis of eDNA. The statistical model for eDNA is presented in the main text.

### Standards

For each qPCR plate, we included known samples with a known number of hake DNA copies to allow us to translate the observed PCR cycle at which amplification occurred to the number of copies in a sample. We present both the relationship for the PCR cycle at which amplification was observed (Fig. 1) and the probability of observing amplification at different DNA copy numbers (Fig. 2). The analysis of standard show that our qPCR analyses were sensitive to very low copy numbers and consistent across plates.

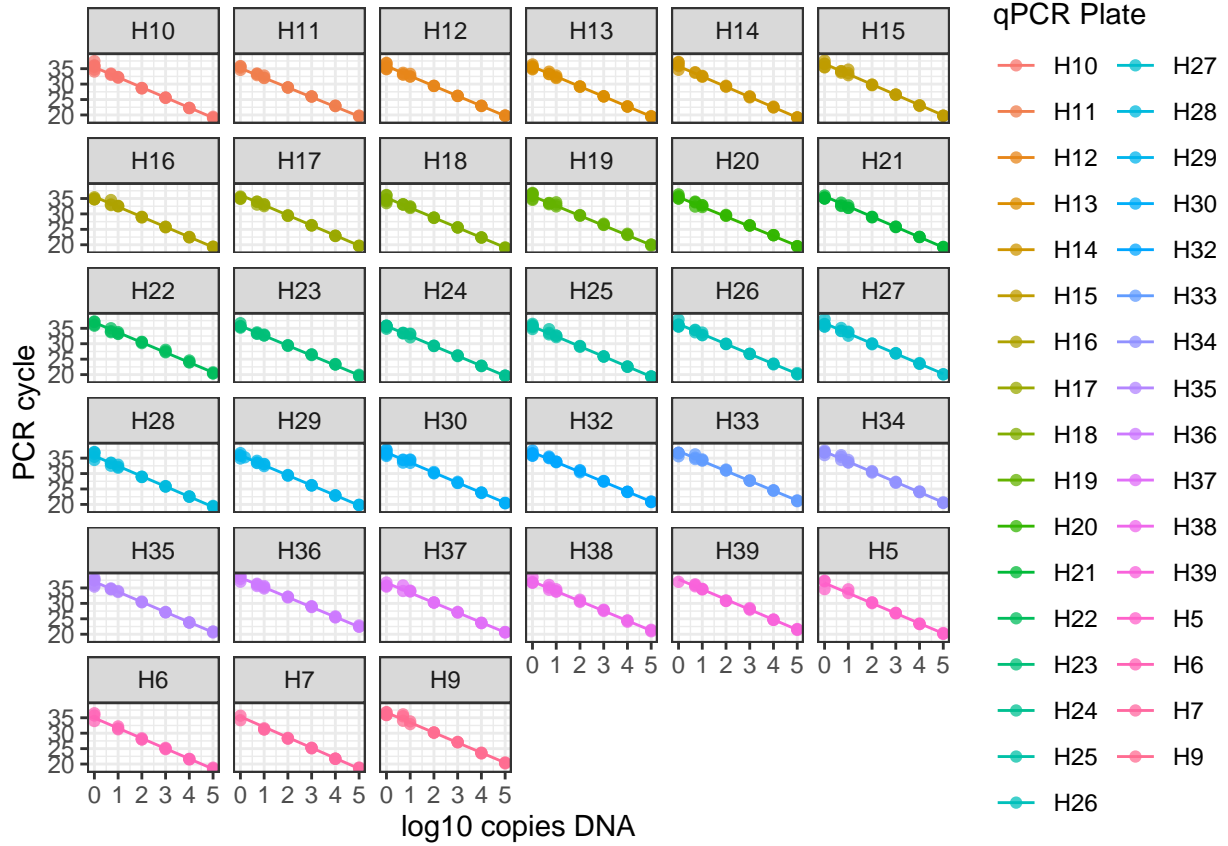


Figure 1: qPCR standards for 33 plates with DNA copy numbers ranging from 1 to 100,000.

### Investigations of contamination.

In addition to field samples, we collected two kinds of control samples **ship board** to look for potential contamination of samples during sample processing. We found low levels of contamination in while control

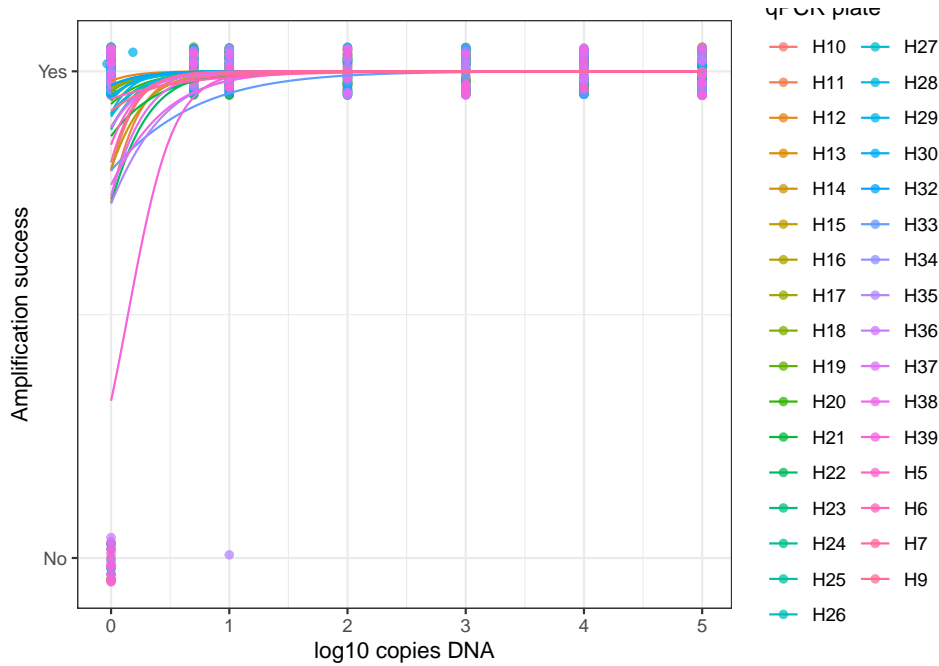


Figure 2: Amplification success for qPCR standards for 33 plates with DNA copy numbers ranging from 1 to 100,000.

samples. While 18 of 49 had hake DNA below the detection limit of  $20 \text{ copies } L^{-1}$ , 31 had higher concentrations. Of these 31 samples, most (21 samples) had very low concentrations  $< 100 \text{ copies } L^{-1}$  but there was a single sample with  $> 500 \text{ copies } L^{-1}$  (Fig. 3). From these observations, we estimated the parameters from a log-normal distribution to describe the distribution of hake DNA from contamination (Fig. 3; posterior mean estimates:  $\mu = 3.498$ ,  $\sigma = 0.769$ ). This corresponds to an expected contribution of 44.4 copies per liter per sample. For this distribution 92% of samples would be expected to have less than  $100 \text{ copies } L^{-1}$  but there will be a small number of more significantly contaminated samples. We remain unsure of the ultimate source of contamination in these samples and are conducting

To understand the potential effect of contamination on our results, we first plotted the estimated distribution from contamination against the estimates of hake DNA concentration in field samples (Fig. 4). The observed hake DNA distributions were all substantially larger than estimated contamination distribution. So while contamination may be playing a minor role in determining the observed hake DNA, it is not the dominant factor determining the spatial or depth specific distribution of DNA. We cannot identify which samples have a significant contribution of contamination but there is no suggestion that the contamination has a spatial or depth-specific effect.

MORE?

## Inhibition

As part of our qPCR protocol (Ramón-Laca, Wells, and Park, n.d.), we included an internal positive control (IPC) of the reaction to account for PCR inhibition. Any sample that had a difference of more than 0.5 PCR cycle from the IPC was considered to be inhibited. Inhibited samples were initially purified using an inhibitor removal column and thereafter diluted 1:2, 1:5 and 1:10 to circumvent inhibition and requantified. We detected inhibition primarily in the near-surface samples (Fig. 5). After inspection of qPCR results following dilution, we included only 1:5 dilution samples in the final statistical model.

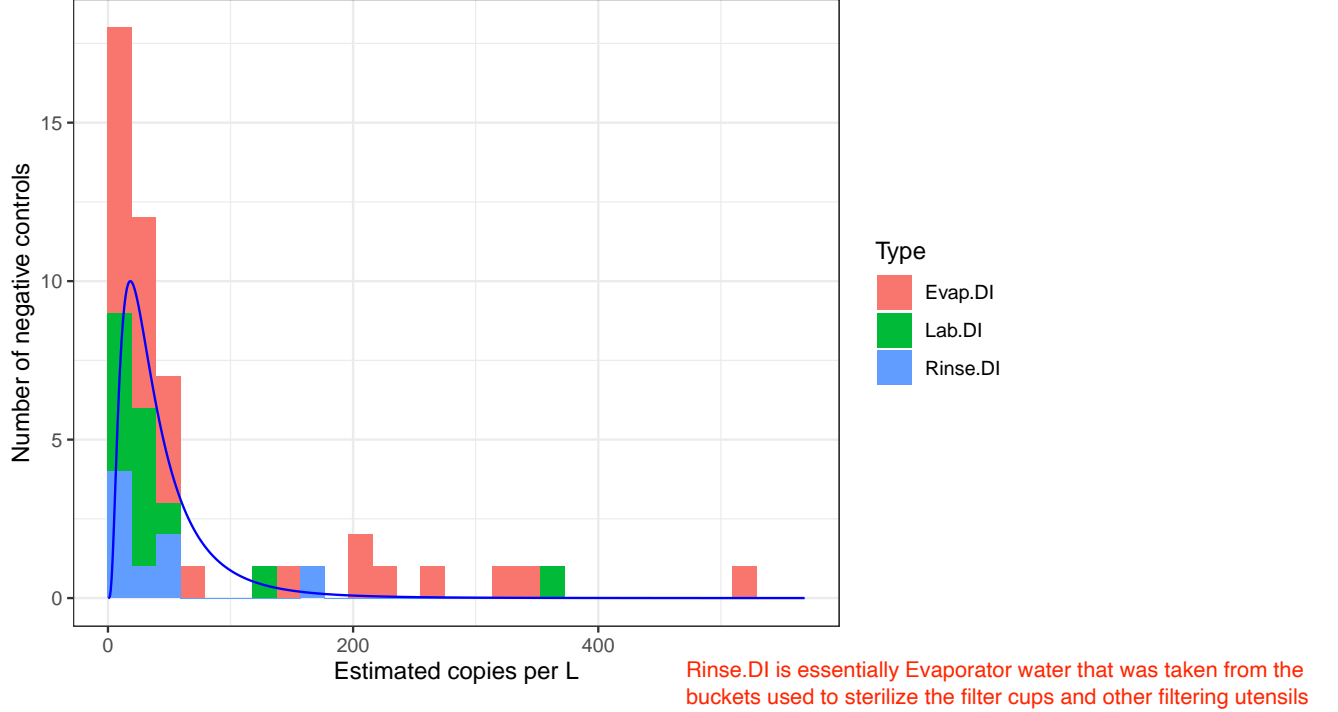


Figure 3: Histogram of point estimates for hake DNA copies per L from the **two types** of field control samples (N=49). Blue line shows estimated lognormal distribution from the observed hake copies per L.

### Ethanol wash error

For a subset of samples, DNA extraction was incorrectly washed with overdiluted ethanol at the final desalting step of the DNA purification. Specifically, samples were washed with a 30% ethanol solution instead of a 70% ethanol solution. Affected samples lost some DNA during the washing processes. Washed bottles were geographically restricted to the northern part of the sample area, but occurred across all sample depths (Fig. 6). To examine the magnitude of this problem, we experimentally divided samples taken from 26 individual station-depth combinations and subjected one bottle taken from those stations to the proper 70% ethanol wash and for the paired bottle used a 30% ethanol wash. This paired design allowed us to estimate the magnitude of the DNA concentration lost due to the wash error.

Within the model we estimate that this washing error led to  $\omega = -0.56$   $[-0.71, -0.42]$  (mean [95% CI]), which translates to approximately 72.73% of the DNA being lost due to the ethanol wash error (Fig. 7).

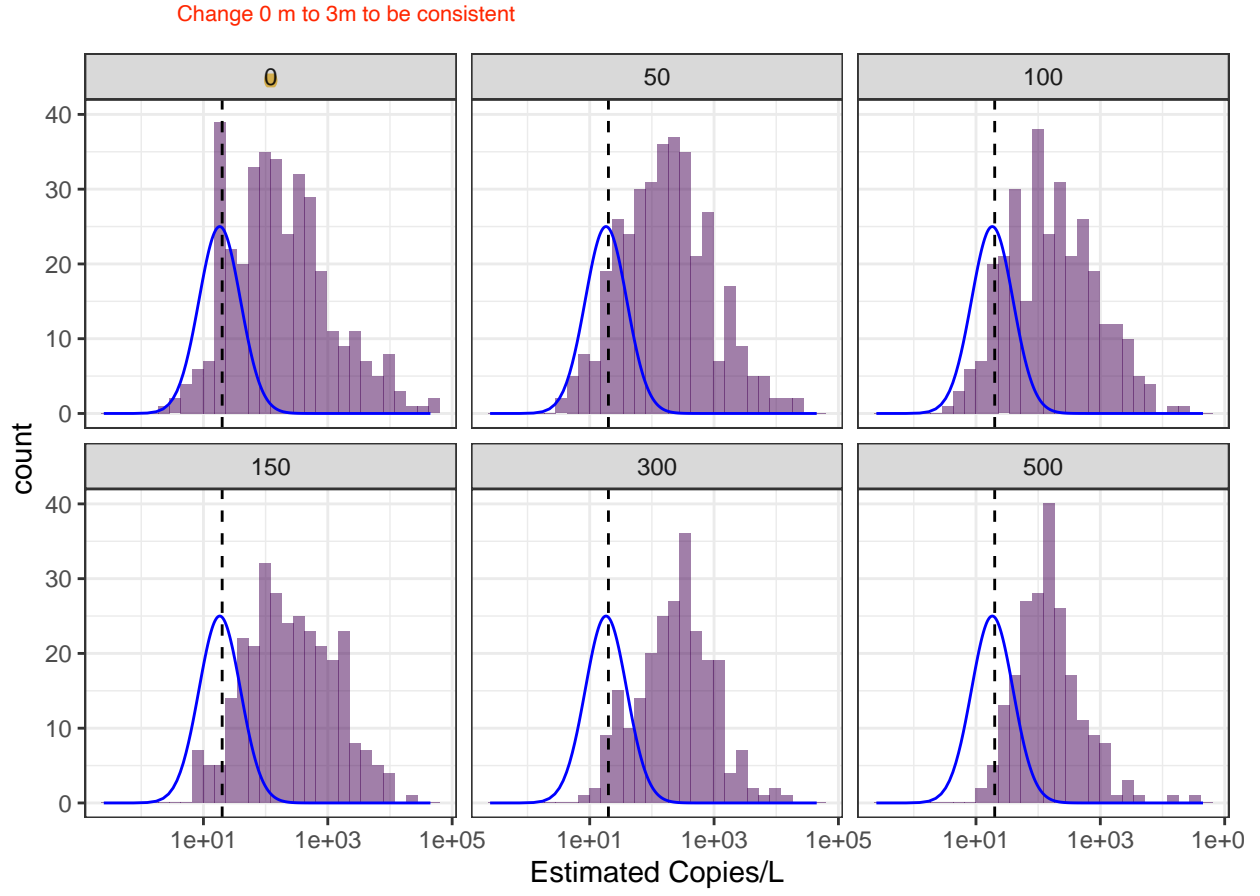


Figure 4: Empirical histograms of estimated hake DNA copies per L for individual samples at six water depths (purple) and estimated distribution of contamination derived from control samples. Vertical dashed line shows 20 copies per L, the putative detection threshold for our methodology. blue line

Change 0 m to 3m to be consistent

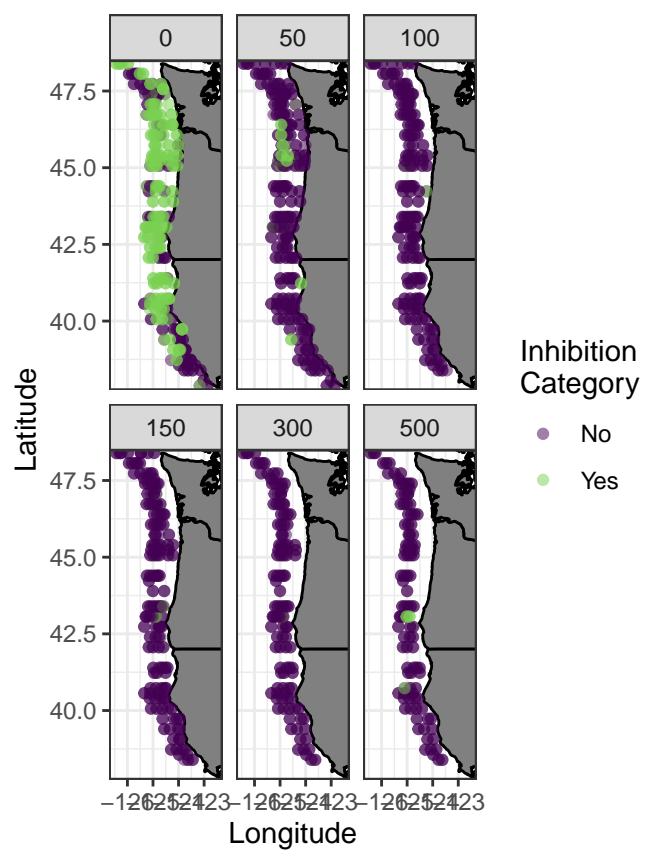


Figure 5: Samples that were identified as being inhibited by sample depths. Samples identified as being inhibited were diluted and re-run using quantitative PCR.

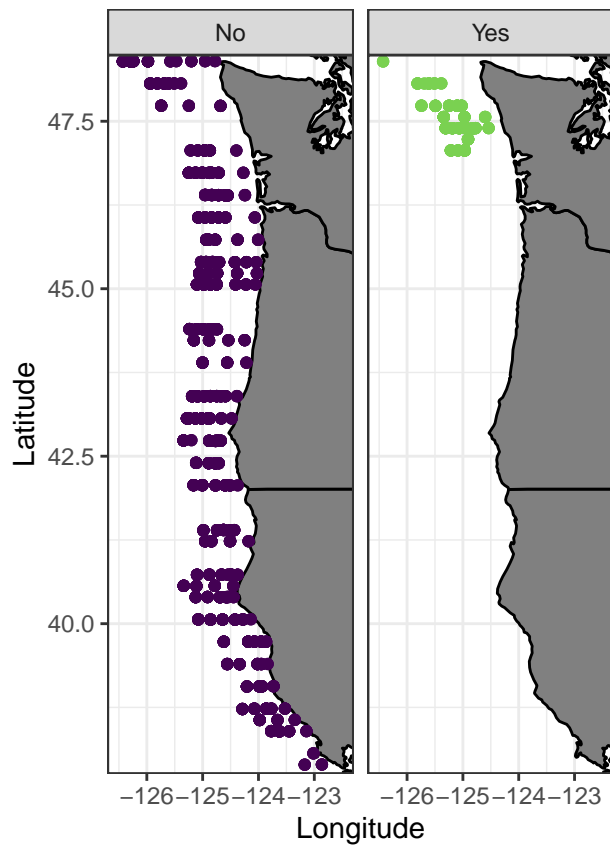


Figure 6: Samples that were identified using the overdiluted ethanol wash. Samples identified as being diluted were included in the statistical model.

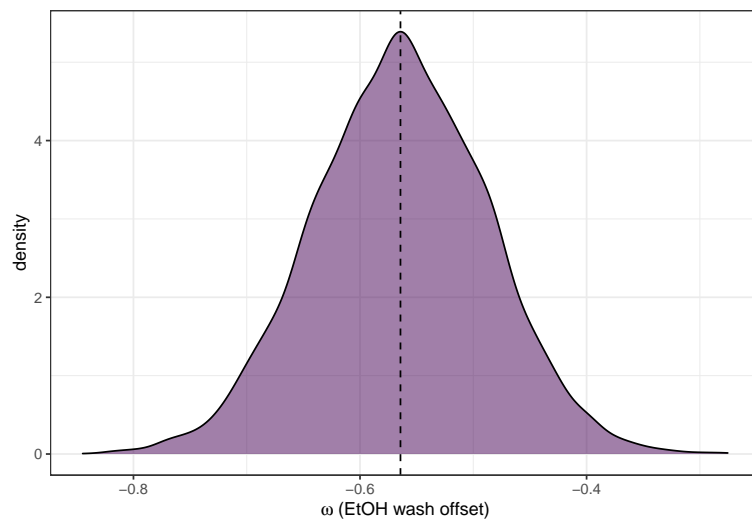
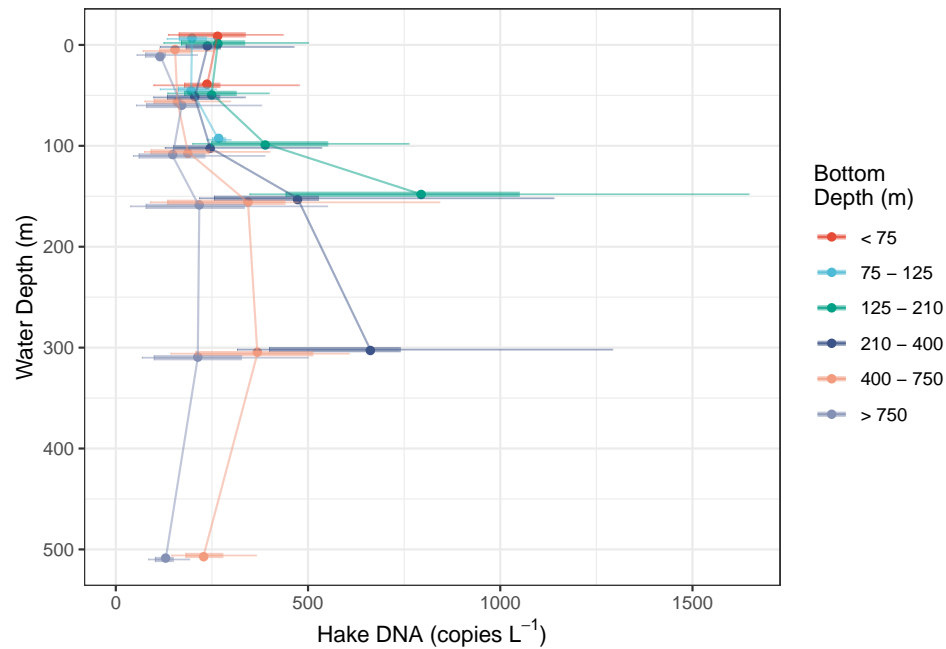


Figure 7: Posterior estimat of OMEGA, the offset due to the ethanol was error. Vertical dashed line indicates posterior mean.

## eDNA model

### Marginal DNA concentration with depth

The maps presented in the main text show spatial variation in DNA concentration it can also be informative to examine generalize patterns of DNA concentration across both bottom and the water depth at sampling. We found strong patterns of DNA concentration varying with depth when averaged across space (Fig. )



Posterior estimates of  $D_{xyd}$  (the hake DNA concentration at each station-depth combination) by the water depth sampled and categories of the depth of the bottom. The distribution of mean  $D_{xyd}$  among station-depths (mean, interquartile range, and 90 CI among station-depths). Bottles at a sample location become increasingly similar at deeper sampling depths

### CV maps for eDNA

Comparison of eDNA and Acoustics (fig. 2 main text) but with 0.5 degree breaks.

Need to add the map just like figure 2 in main text.

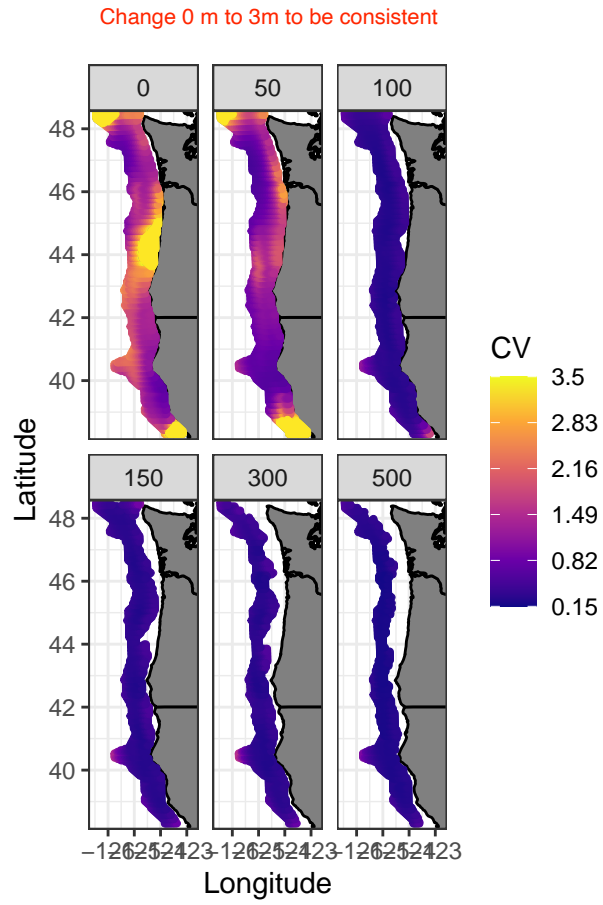


Figure 8: Posterior estimates of the CV of hake DNA concentration by sampled depth



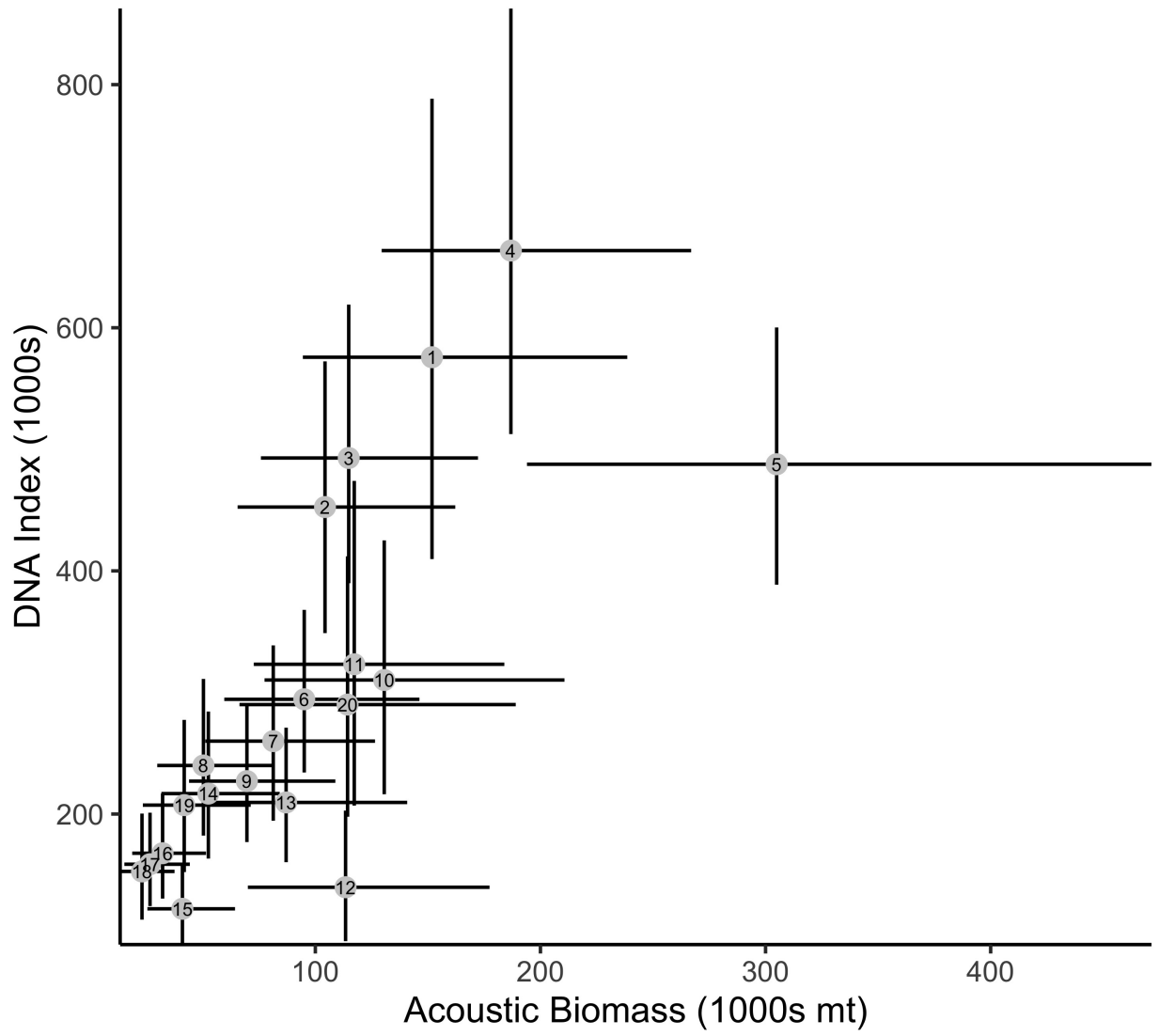


Figure 9: Pairwise comparison between DNA and acoustics-derived biomass. Correlation between methods among the 20, half degree bins (posterior mean, 90 CI;  $RHO = 0.74$ ). Numbers indicate associated region identified in Fig. 2.

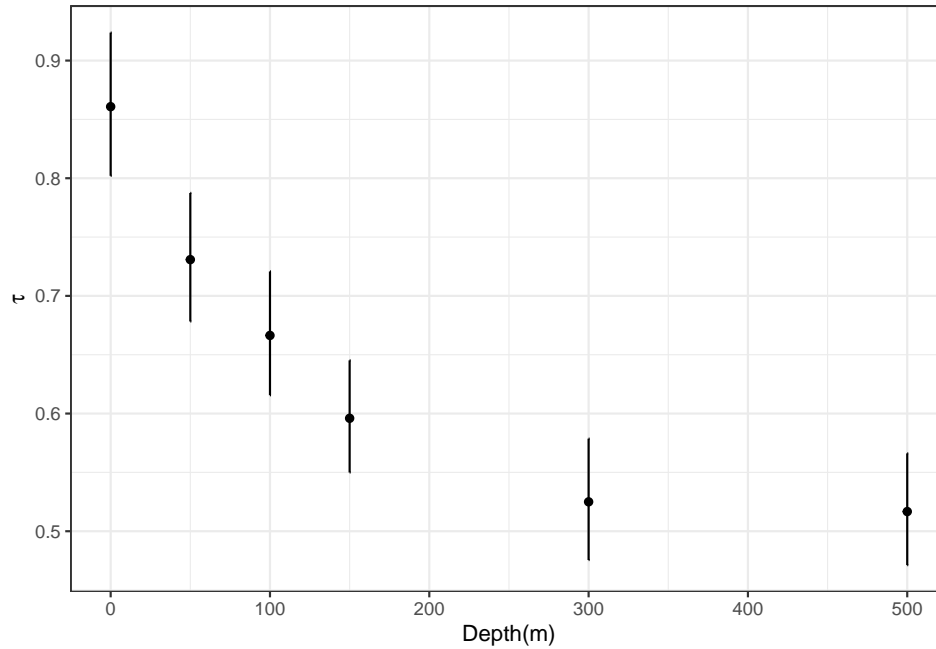
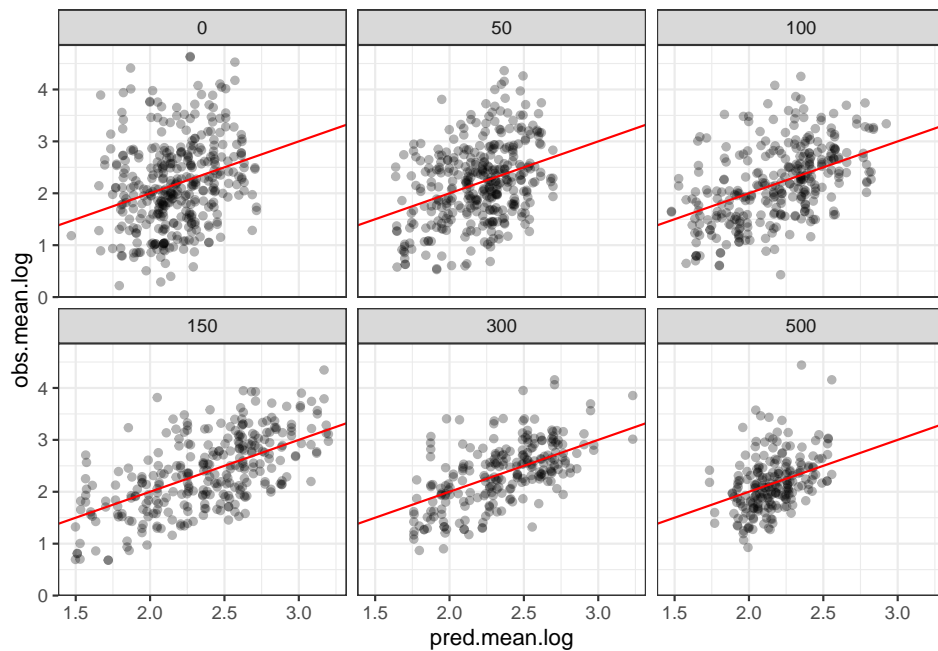


Figure 10: Estimates of TAU (log10 standard deviation of hake eDNA concentration between bottles at a single station-depth) for each sampling depth (posterior mean and 90 CI shown). Bottles at a sample location become increasingly similar at deeper sampling depths

## Sources of Uncertainty

Raw predicted-observed plots for both models.



\begin{figure}  
Predicted vs. observed plots by sample depth. D<sub>xyd</sub> (predicted mean) vs D<sub>xydi</sub> (observed in each Niskin bottle). \end{figure}

Add plot of Eta and Sigma.

**Table of prior distributions for eDNA**

**Acoustic-Trawl model**

**Table of prior distributions for acoustic-trawl suvey**

**Components of the acoustic model - pres-abs, positive, unconditional**

**Pred - Obs plots**

**Posterior variance for acoustics**

**Citations**

Ramón-Laca, Ana, Abigail Wells, and Linda Park. n.d. “A Workflow for the Relative Quantification of Multiple Fish Species from Oceanic Water Samples Using Environmental Dna (eDNA) to Support Large-Scale Fishery Surveys.” *PLoS NA*: NA.