# Bottom temperature at CSI1

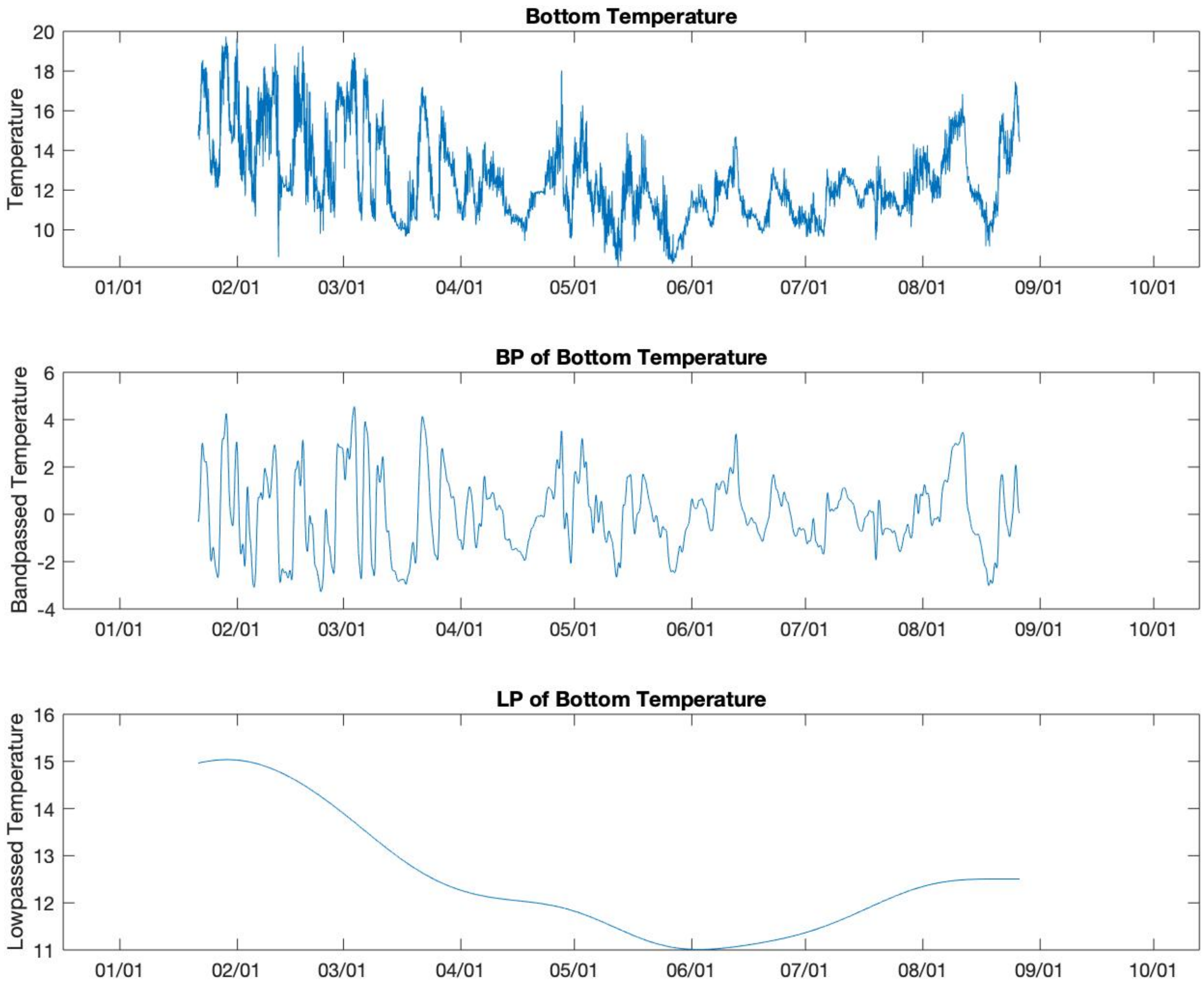


Figure 1: The top plot above shows the bottom temperature for the first CSI deployment. In the middle plot, this time series is bandpass filtered to show primarily the fluctuations due to the Gulf Stream front moving back and forth. The bandpass filter removes higher-frequency tidal fluctuations as well as the lower-frequency or seasonal fluctuation. The long-term variability is captured in the bottom plot, a lowpass filtered version.

# Time-series comparisons (using bottom temperature)

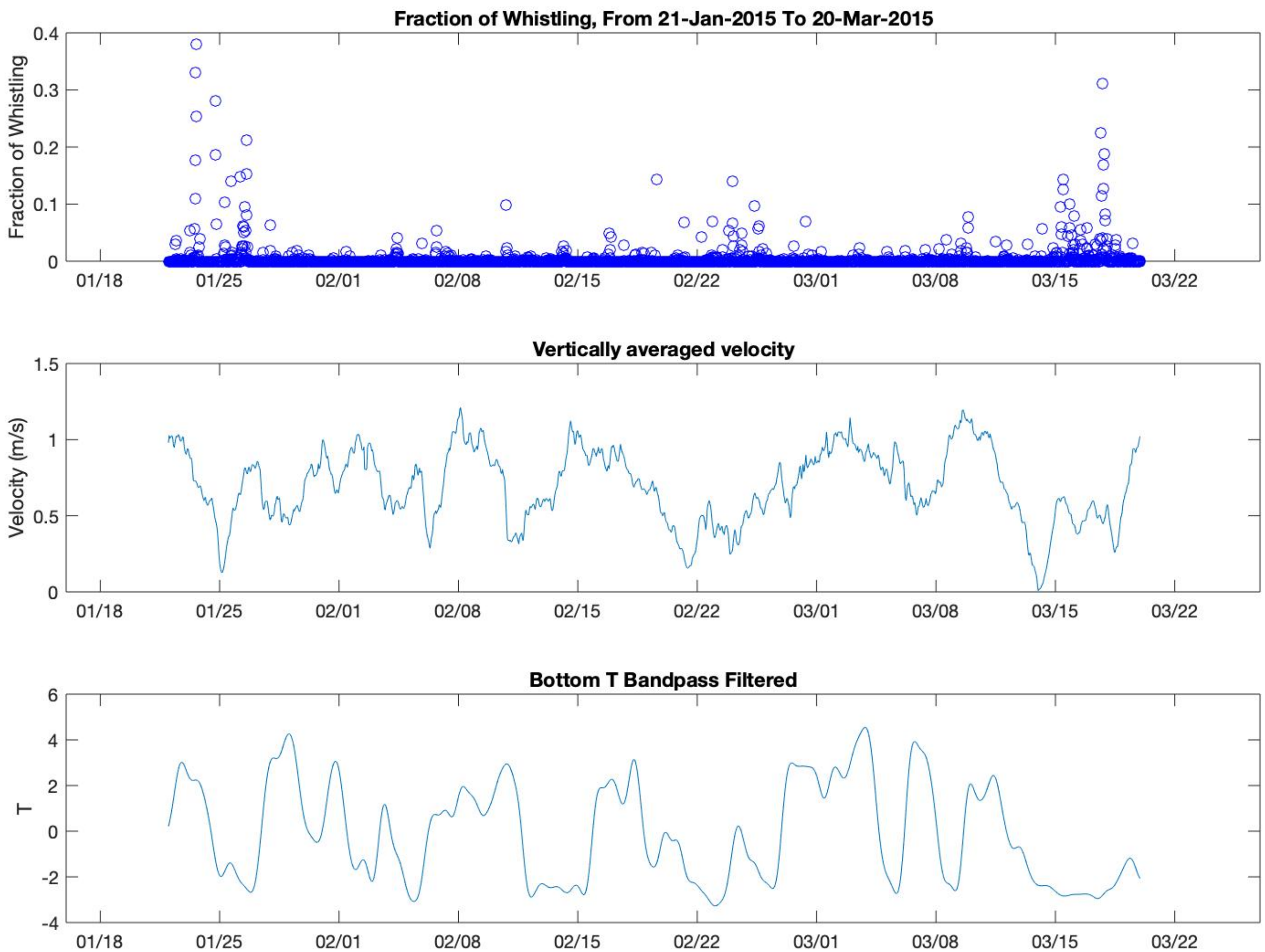


Figure 2: In these three plots, I zoom in on the Winter of 2015, during the first CSI deployment. In the top plot is a marine mammal vocalization metric—the total number of seconds of detected whistles in a recording interval divided by the duration of the recording interval. In the middle plot, from the ADCP, I plot the vertically averaged along-stream velocity. In the bottom plot, the bandpass filtered bottom temperature. When the Gulf Stream is meandering away from the mooring site, the velocity above the mooring is low and the bottom temperature is cooler i.e. the bandpass filtered fluctuation is negative. In this time period, if there is a lot of whistling, it is likely that the Gulf Stream is meandering away from the mooring site.

# Segment the temperature data based upon vocalization metric

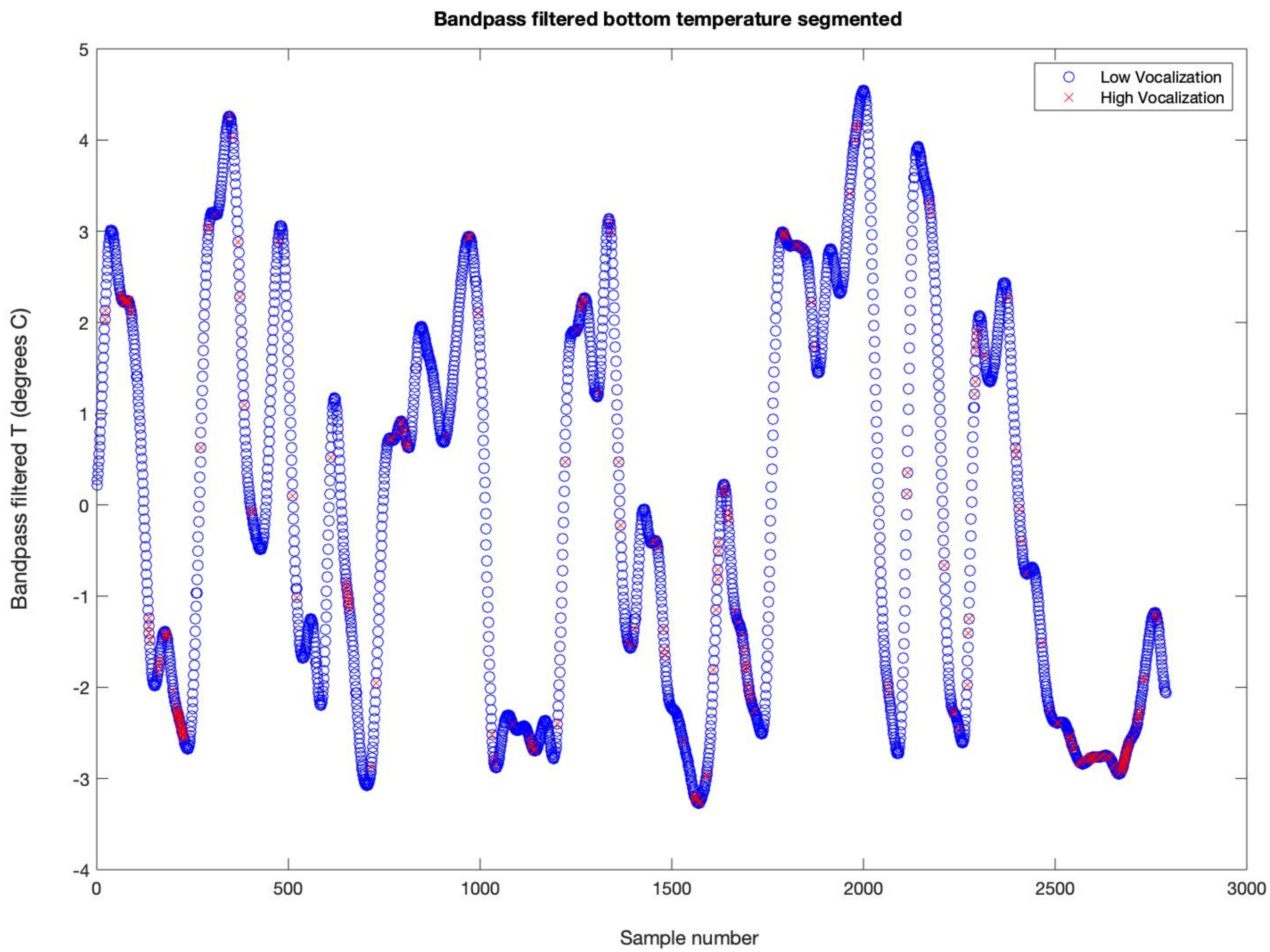


Figure 3: The hypothesis is that the marine mammals favor the cooler water in the meander trough. To test this hypothesis, we segment the temperature data into two subsets. The first ("high vocalization") subset $T\_BP_{HV}$ includes the temperature samples that were acquired when there was "a lot of whistling". The second ("low vocalization") subset $T\_BP_{LV}$ includes all other temperature samples. By "a lot of whistling", I mean: the recording intervals where the vocalization metric was in the top 10%.  This figure shows this segmentation applied to the Winter of 2015.

# Compare PDF and CDF of different segments

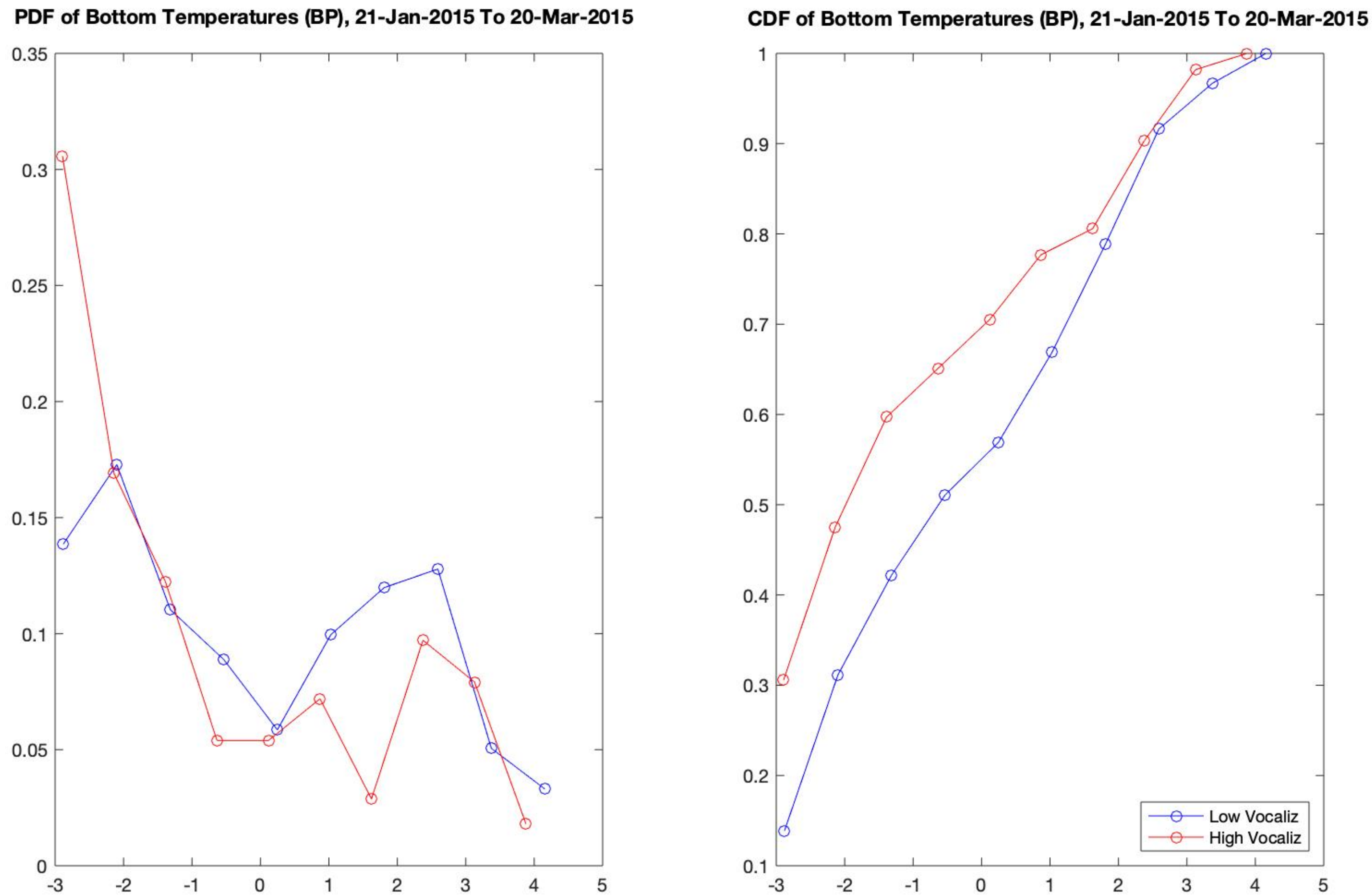


Figure 4: To test the hypothesis, we can see if the data in the two segments are distributed differently. For the winter of 2015, the figure on the left compares the probability density functions (PDFs) of the two segments. The plot on the right compares the cumulative distribution functions. These plots show that—for the winter of 2015—the high-vocalization segment is skewed towards the cooler water.

## Use the rank-sum test to compare segments per season (for bottom temperature)

To quantify the difference between the two segments ($T\_BP_{LV}$ and $T\_BP_{HV}$), we create the following features:

- The difference between the median values:

  $$diff\_Medians = median(T\_BP_{HV}) - median(T\_BP_{LV})$$

- The p-value from the rank-sum test:

  $$p\text{-}value = ranksum(T\_BP_{LV}, T\_BP_{HV})$$

Again, our hypothesis is that the water is cooler when there are more vocalizations. This would be the case if:

- diff_Medians is negative (indicating cooler water in the high-vocalization set)
- | diff_Medians | is large, indicating the two segments are different
- p-value is very small, indicating the two segments are statistically different with high confidence

In Figure 5, these features are plotted on a scatter plot. This is done per season for three deployments (CSI1, CSI2, ADEON1).

## Use the rank-sum test to compare segments per season (for bottom temperature)

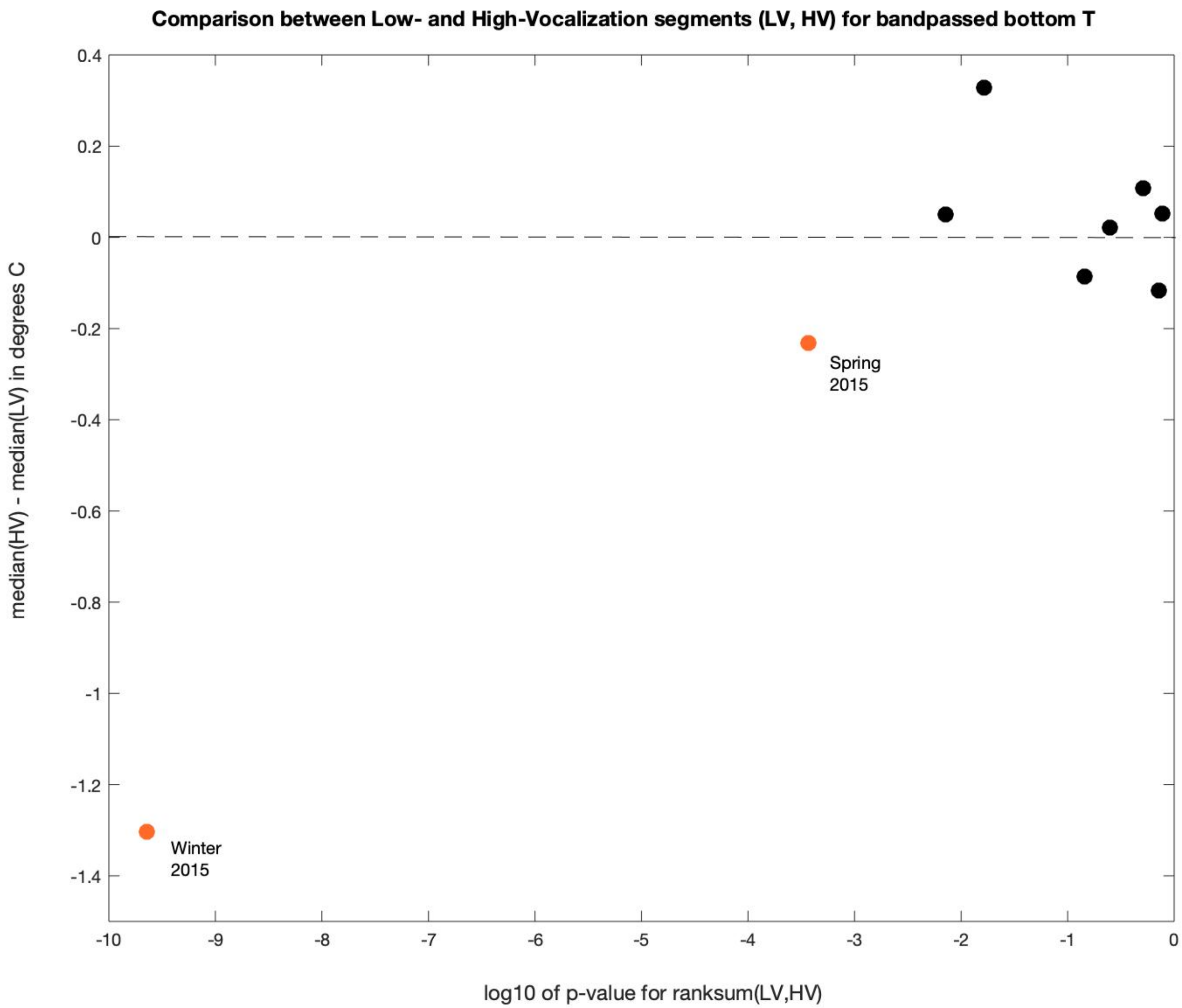


Figure 5: Each point represents a season. On the x-axis is the log10(p-value) from the rank-sum test. On the y-axis is diff_Medians. Winter and Spring of 2015 are color-coded orange to highlight the fact that our hypothesis is confirmed for these seasons. During Winter of 2015 (and to a lesser extent Spring of 2015), the high-vocalization segment was cooler AND statistically different than the low vocalization segment. One could not say the same for the other seasons (in black)—where either the p-value is too high or the difference between the medians is not negative.

# Time-series comparisons (using SST)

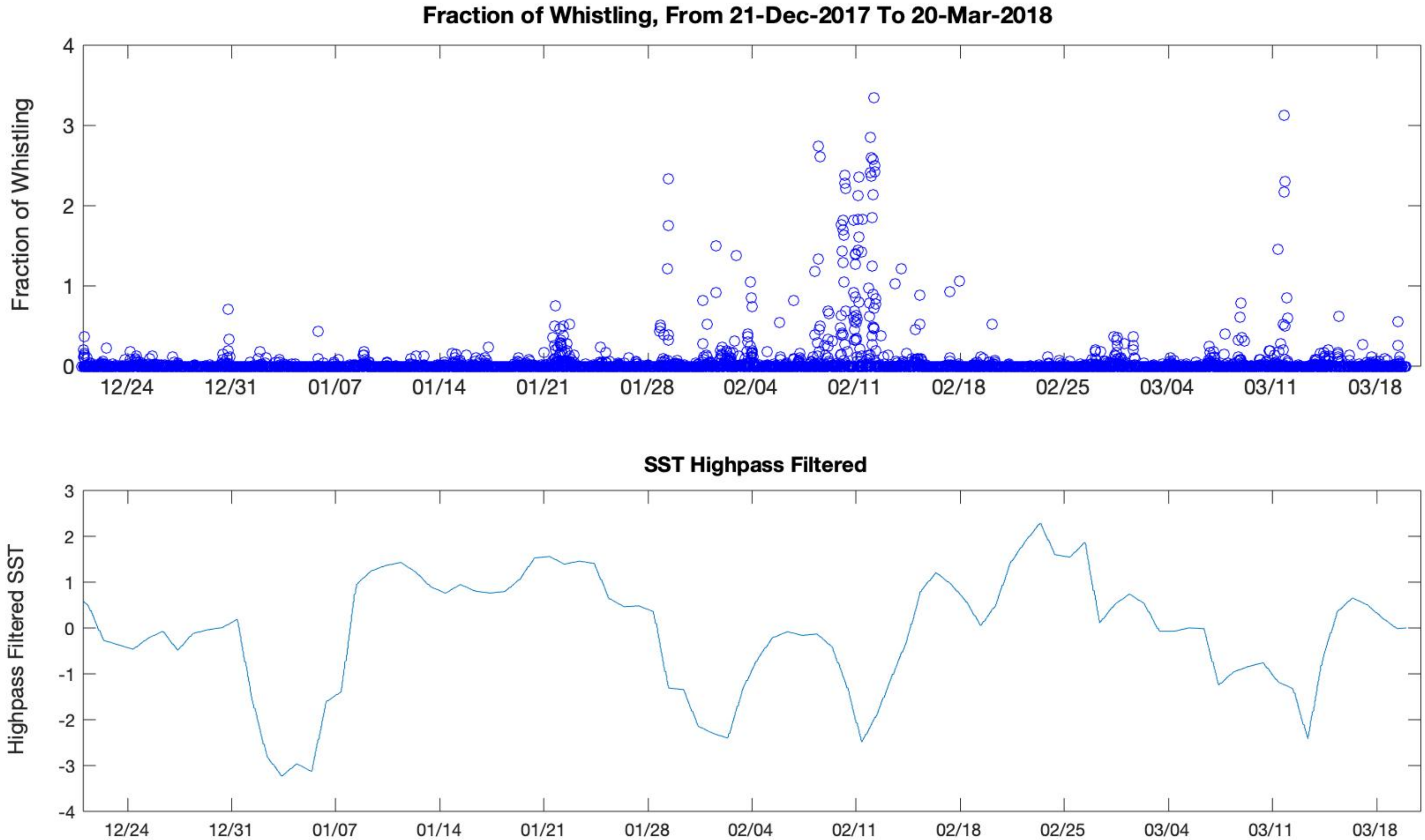


Figure 6: In these two plots, I zoom in on the Winter of 2018, during the ADEON EN615 HAT deployment. In the top plot is a marine mammal vocalization metric—the total number of seconds of detected whistles in a recording interval divided by the duration of the recording interval. In the bottom plot, the highpass filtered sea surface temperature (SST) above the mooring. Periods with a lot of whistling occur when the SST is lower.

# Use the rank-sum test to compare segments per season (for SST)

Now, we compare vocalizations to highness filtered SST. As with bottom temperature, we create two segments, $SST\_HP_{LV}$ and $SST\_HP_{HV}$ .

To quantify the difference between the two segments ($SST\_HP_{LV}$ and $SST\_HP_{HV}$), we create the following features:

- The difference between the median values:

  $$diff\_Medians = median(SST\_HP_{HV}) - median(SST\_HP_{LV})$$

- The p-value from the rank-sum test:

  $$p\text{-}value = ranksum(SST\_HP_{LV}, SST\_HP_{HV})$$

Again, our hypothesis is that the water is cooler when there are more vocalizations. This would be the case if:

- diff_Medians is negative (indicating cooler water in the high-vocalization set)
- | diff_Medians | is large, indicating the two segments are different
- p-value is very small, indicating the two segments are statistically different with high confidence

In Figure 7, these features are plotted on a scatter plot. This is done per season for three deployments (CSI1, CSI2, ADEON1).

In Figure 7, our hypothesis was confirmed for ADEON's Winter of 2018 (using SST). In Figure 5, our hypothesis was confirmed for CSI's Winter of 2015 (using bottom temperature).

# Use the rank-sum test to compare segments per season (for SST)

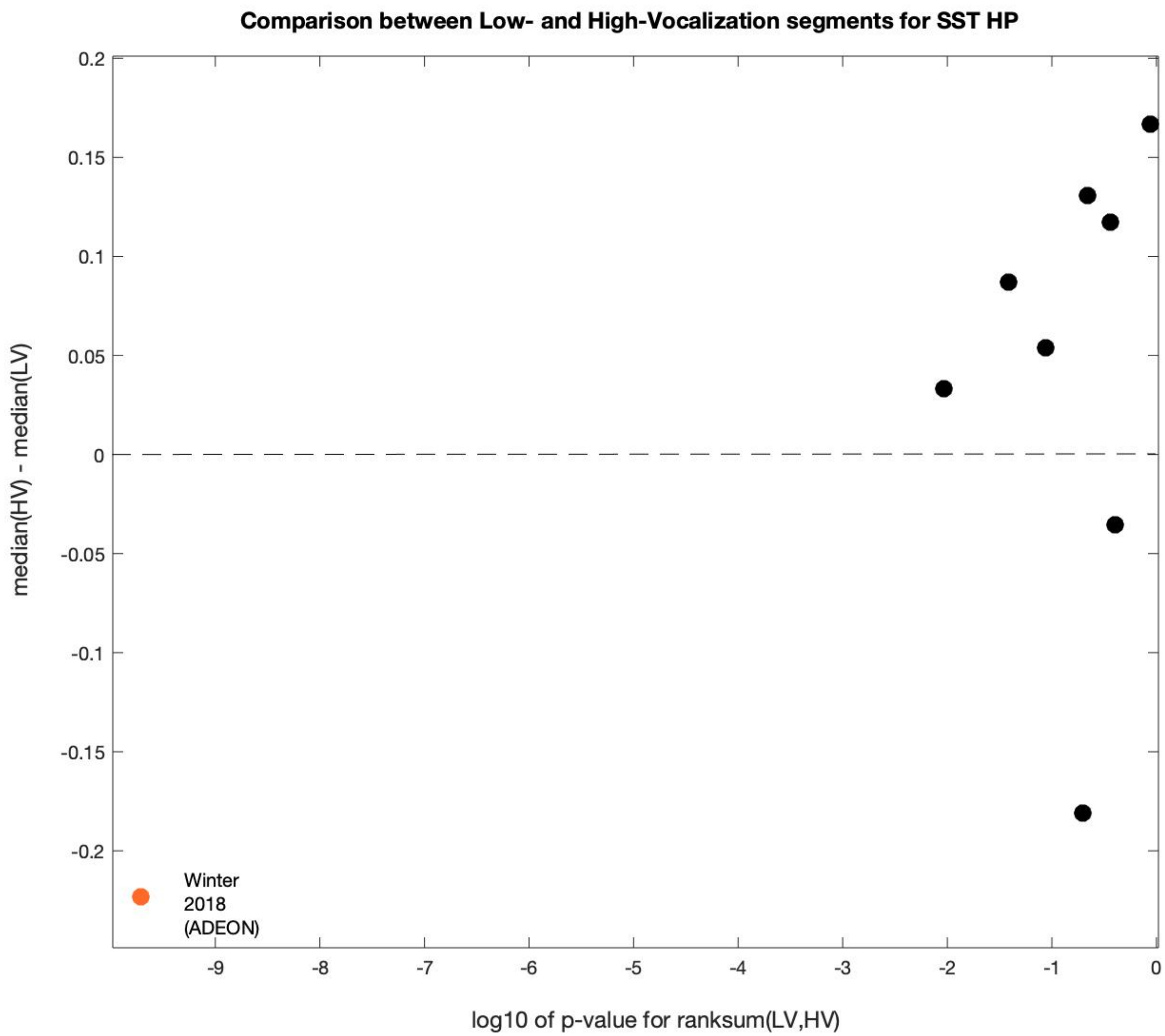


Figure 7: Each point represents a season. On the x-axis is the log10(p-value) from the rank-sum test. On the y-axis is diff_Medians. Winter of 2018 (ADEON) is color-coded orange to highlight the fact that our hypothesis is confirmed for this season. During Winter of 2018, the high-vocalization segment was cooler AND statistically different than the low vocalization segment. One could not say the same for the other seasons (in black)—where either the p-value is too high or the difference between the medians is not negative.