Season 2: exploratory analysis

July 31, 2018

Density & Detection Estimates

We estimated the density in each lake following three different survey methods (Figure 1). Distance survey (*in which the distance from the transect line is recorded*), double observer survey (*in which a narrower transect is surveyed and distance is not recorded*), and quadrat survey (*in which a number of smaller quadrats on each transect are surveyed, assuming detection is perfect*). The bars in Figure 1 correspond to one standard error. A potentially interesting outcome is that quadrat surveys are consistenlty estimating higher densities than the other designs, though it is clear from the amount of uncertainty in the estimates that we cannot make very strong conclusions on this point yet. All counts were modeled using design-based estimates.

There are a number of ways to estimate the variance in the surveyed data. We tested three on Lake Burgan below. The total variance in the double observer methods is given by $\text{Var}(\hat{D}) = \hat{D}^2 \left(\frac{\text{Var}(n)}{n^2} + \frac{\text{Var}(\text{E(s)})}{\text{E}(s)^2} + \frac{\text{Var}(a\hat{P})}{(a\hat{P})^2} \right)$. For a design-based estimate with K transects each of length l_k the first variance component is $\text{Var}(n) = \frac{K}{K-1} \sum_{k=0}^{K} l_k^2 \left(\frac{n_k}{t_k l_k} - \frac{\sum n_k}{\sum l_k} \right)^2$. Finally, we applied a model-based estimation process where transect counts were modeled as a negative binomial random variable. We used an offset in the model was the transect survey area times the estimated probability of detection in the transect.

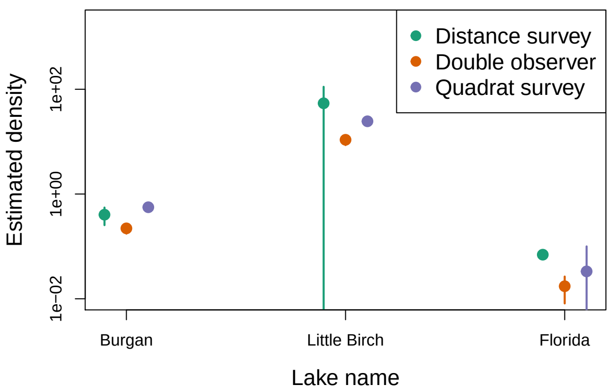
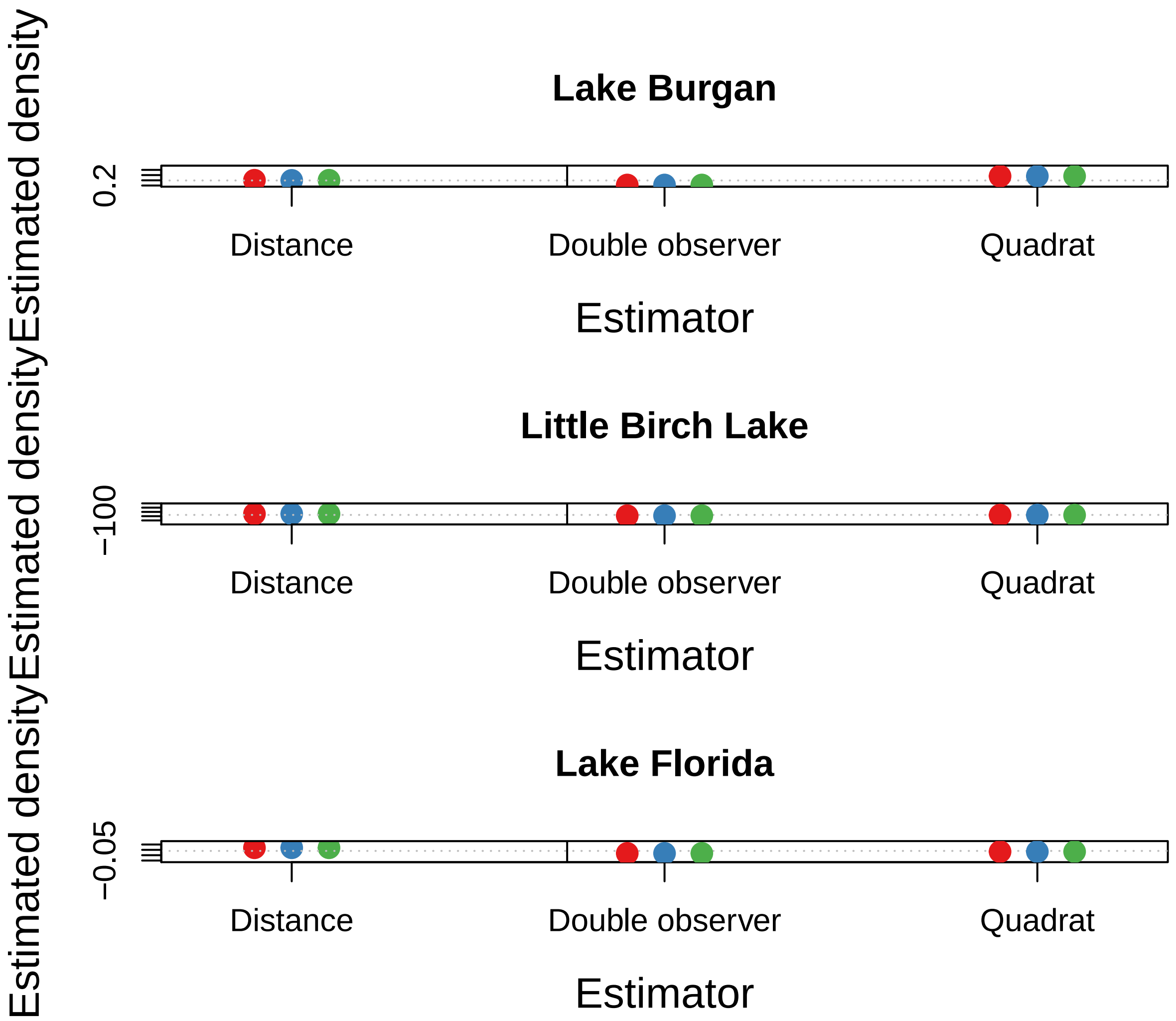


Figure 1: Y-axis is logarithmic. Bars are 2 standard errors.



We can also look at the detection probabilities in each transect (Figure 2). The detection probabilities of the distance survey are consistently lower than the double observer survey. The numbers are the pooled (total) number of detections by both observers. Note that I did not include observer effects in the models (both observers are assumed to have same detection probability).

It's not exactly clear why the double observer survey has such a higher detection than the distance survey, though the estimates are remarkably consistent for each survey type, regardless of sample size. Looking at the histogram of detections (along with the fitted detection model) may provide a clue (Figure 3). I used the hazard-rate detection function, which has a shoulder after which detection falls off.

Figure 3 in Burgan and Florida suggests that detection may fall off at about 0.5 meters. Data is extremely limited to infer any of these patterns, however. Little Birch Lake, which has the most detections, does not show this pattern…

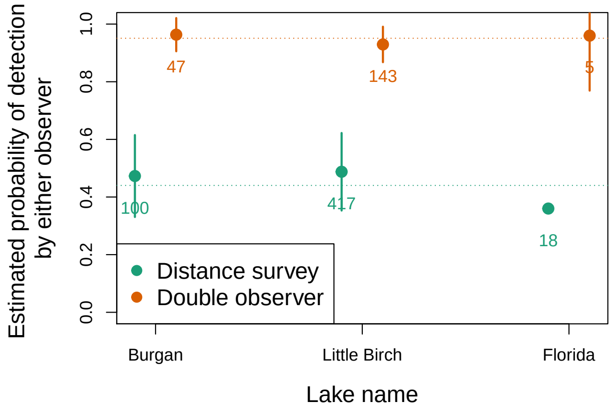


Figure 2: Numbers on the figure denote the total number of detections used to estimate detection probabilities. Bars are two standard errors. Horizontal dashed lines give the average of the lake estimates.

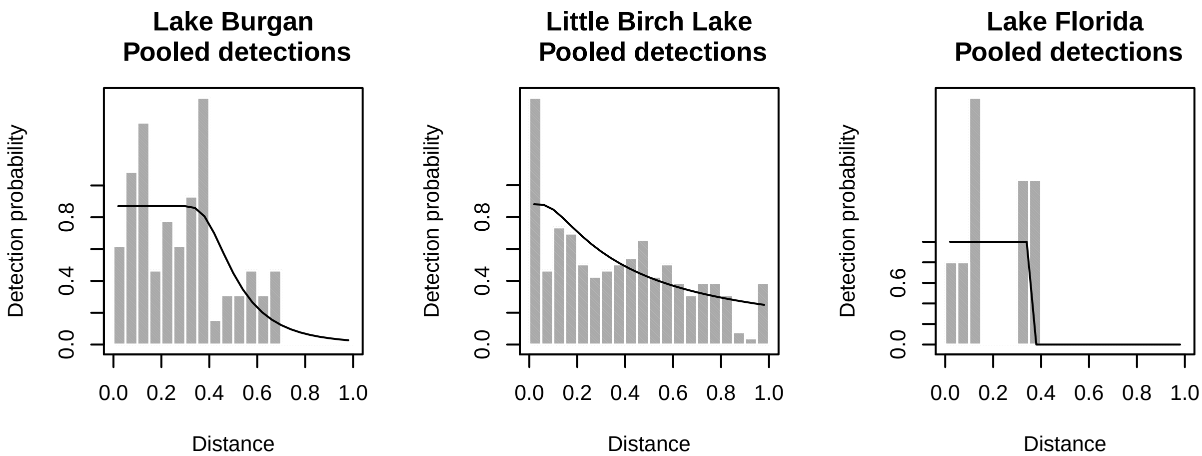


Figure 3:

Table 1: Number of detections by primary and secondary observers. In the quadrat survey only one observer is used.

	Distance survey		Double observer survey		Quadrat survey
Lake name	Primary	Secondary	Primary	Secondary	Detections
Lake Burgan	39	61	38	47	40
Little Birch Lake	165	252	105	143	526
Lake Florida	9	9	4	5	2

Table 2: Correlation between phase 2 detection rates and phase 3 densities estimated using Distance sampling, Double observer sampling, and Quadrat sampling.

	Distance	Double	Quadrat
Burgan	0.29	0.33	0.42
Little Birch	0.63	0.94	0.59
Florida	-0.27	-0.22	0.00

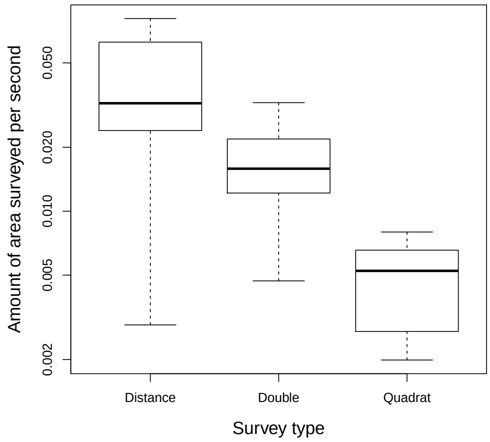


Figure 4: Add color, center

Finally, we can look to see if the Phase 2 surveys, where we went out perform a timed qualitiative search for zebra mussels, are correlated with the densities from phase 3. Table 2 compares the detection rates in phase 2 with the observed densities from phase 3. It looks like there might be some predictive value, but it may be dependent on density and potentially on survey type. Just looking at correlations suggests that the phase 2 count-rate is most strongly correlated with quadrat densities.

Time Budget Summaries

We are also interested in how long it takes to do each task associated with the zebra mussel surveys. Figure 4 is the overall amount of time spent doing a given type of survey versus the amount of area that was actually covered.

We see that the distance sampling is more efficient overall at covering area that the other types of survey.

We can also look at the time spent on the setup, taking habitat data, and on making and recording the encounters. This will allow us to better understand what is driving the relatively

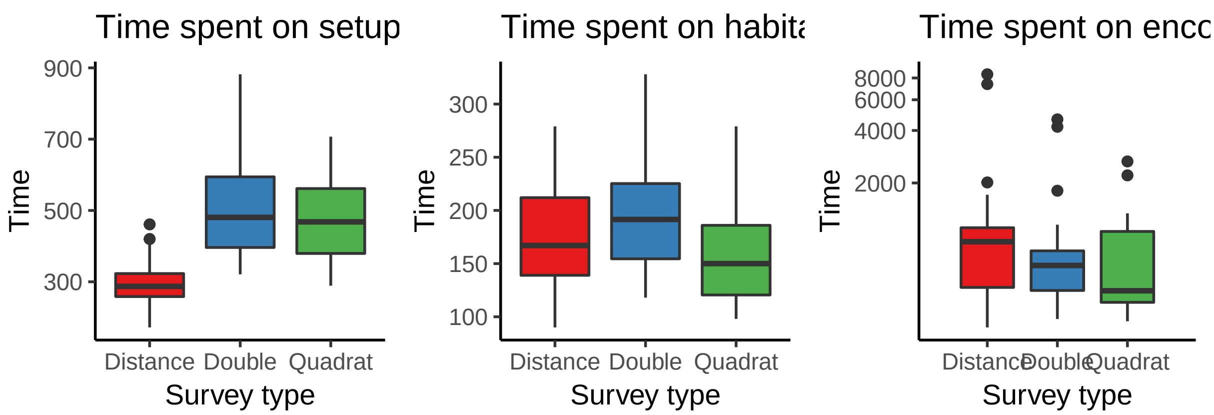


Figure 5:

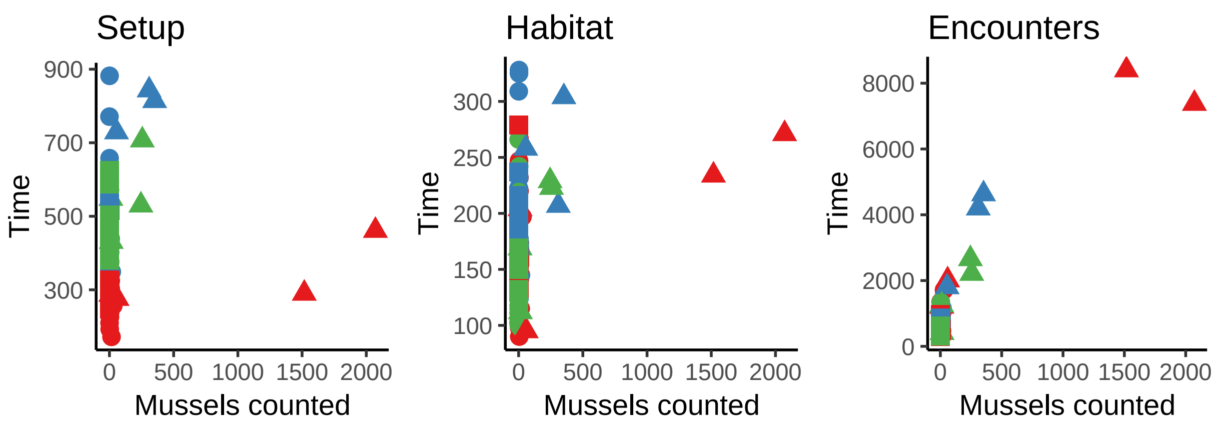


Figure 6:

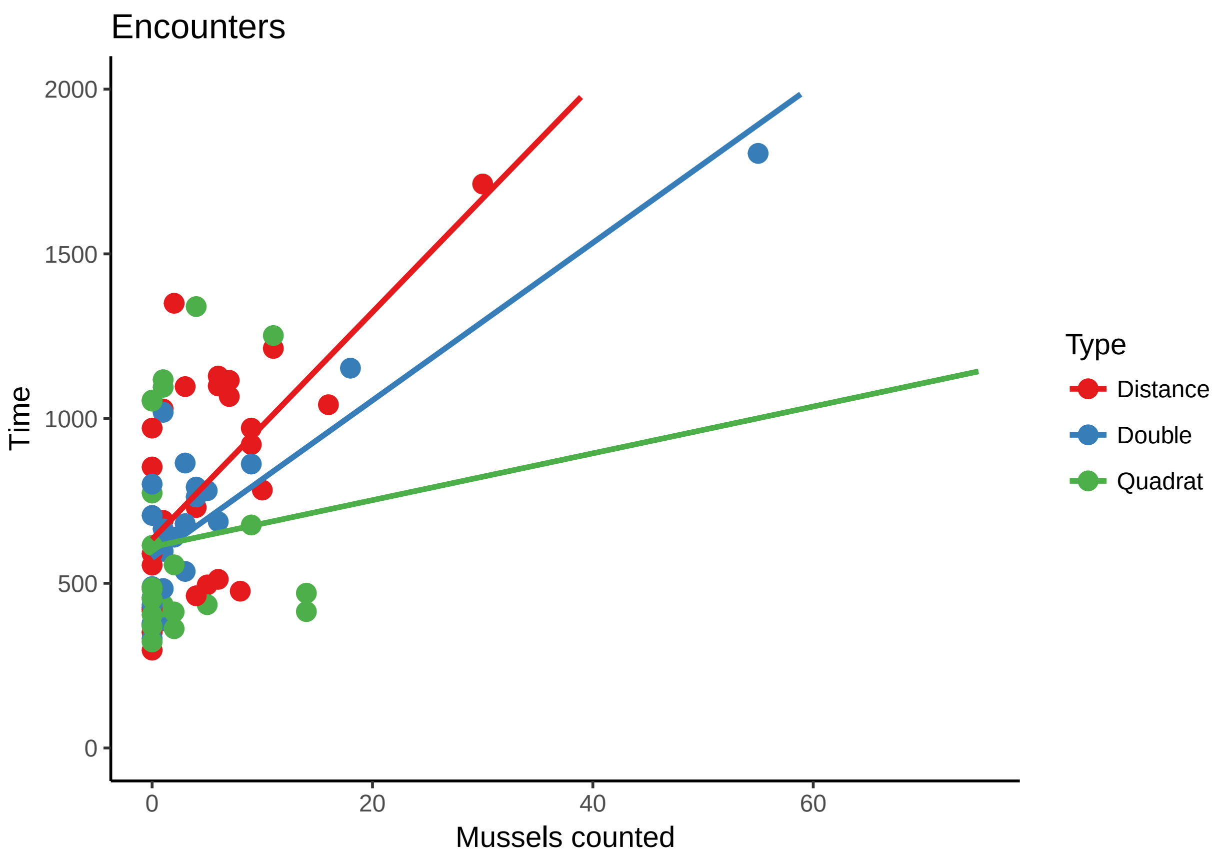
high efficiency of the distance surveys (Figure 5).

It appears that the economy of distance surveys is coming from the setup (laying out transect lines and taking transect-level data) of the transects. This is consistent with reports from the divers. We assumed that the amount of time spent on encounters in the double observer survey would be less than the distance survey, though Figure 5 suggests that any differences are small enough not to matter in the overall efficiency. Note that these times are not standardized to transect lengths at all...

```
## Warning: Removed 7 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 7 rows containing missing values (geom_point).
```

```
## Warning: Removed 55 rows containing missing values (geom_smooth).
```


Encounters
Time
2000
1500
1000
500
0
0
20
40
60
Mussels counted
Type
Distance
Double
Quadrat