Body Size Data Processing

NEON collects macroinvertebrate data via fixed-area samplers (e.g., Surber samples) and measures insect body lengths to the nearest mm. While the samplers vary, all mesh sizes are the same (243 um). Fish are collected using 3-pass removal electrofishing within stream reaches that vary in area across study sites. For each collection, the first 50 fish per taxon are measured for total length in mm (Monahan et al. 2020). Thus, the fish data consist of two subsets of data, one for counts of the total number of fish per pass per reach, and another with the length of the first 50 fish caught per taxon per reach (but not per pass). To convert both types of samples into a single format containing abundance (per square meter) of each dry mass, whether macroinvertebrate or fish, multiple data steps were required as described below.

*Macroinvertebrate*

Macroinvertebrate lengths ranged from 1 to 86 mm. However, as shown below, the samplers appeared to undercount insects less than ~3 mm (Figure X). Insects of 3 mm or less tend to have head widths that are less than 250 um (Stoffels et al. 2003), suggesting the possibility that insects of this size or smaller could pass through the mesh. We assumed that this might generate undercounts for insects as seen in Figure X and removed lengths less than 3mm from the analysis.

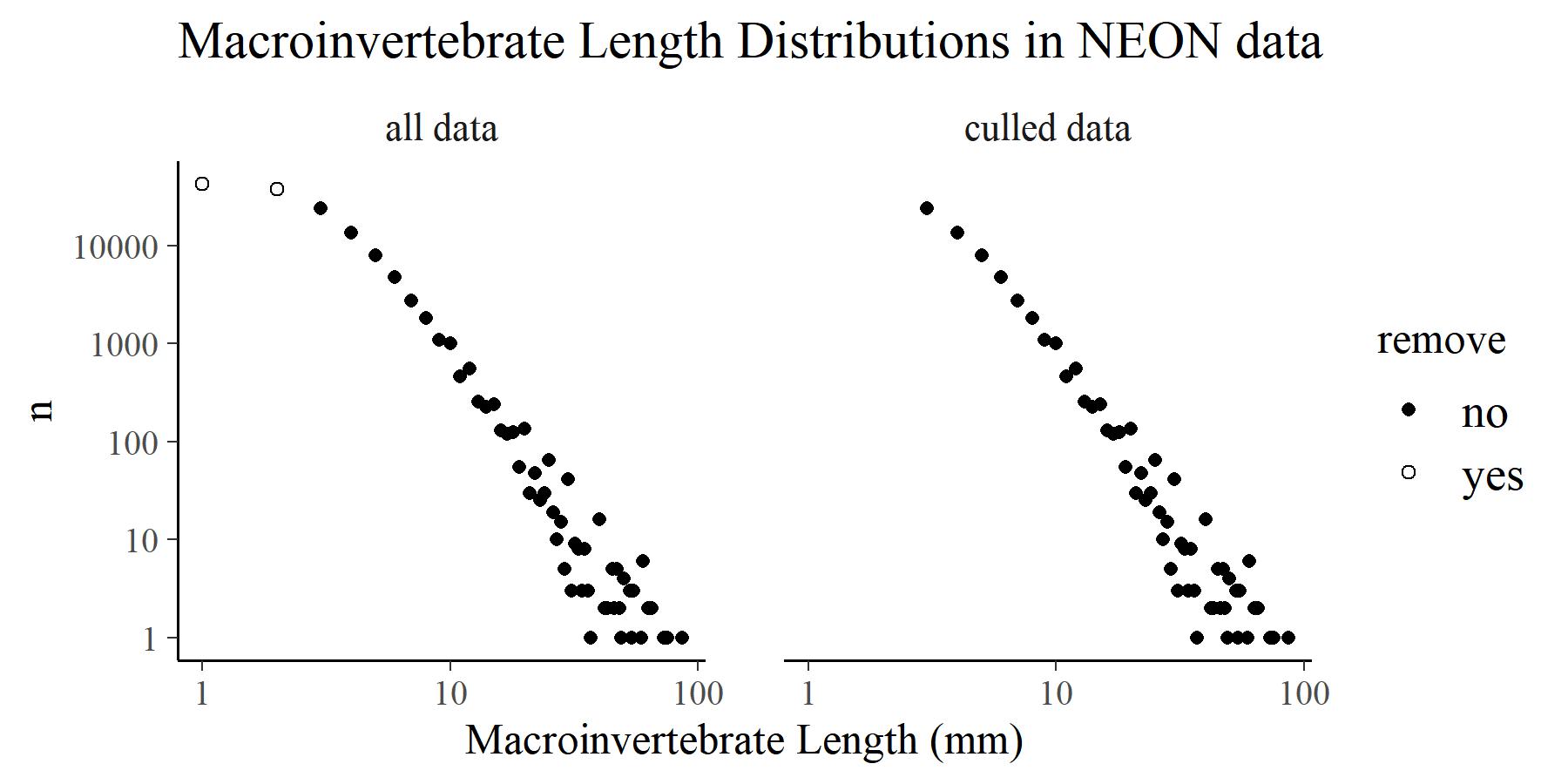


Figure X. The total number (n) of macroinvertebrates in each length class. Data are plotted for all measures in the neon macroinvertebrate dataset totaling 141,246 length measures.

*Fish*

We obtained measurements of individual fish total lengths (mm) from the *fsh\_perFish* table in data product DP1.20107.001 (NEON 2022). For each collection event (i.e., collection date, reach, and site), *fsh\_perFish* contains up to 50 individual total length measurements per species. Most species have fewer than 50 measurements, and in some cases there are more than 50. Length measurements were converted to dry mass estimates (mg) using published length-weight equations specific to the fish species or family (FISHBASE).

There was clear bias in the dataset that suggested undercounting of small fish (Figure X). NEON uses block nets and dip nets that typically have a mesh size of 3.175 mm. Following Garcia et al. (1994), we removed fish with body diameters less than two times the mesh diameter before analyzing the data. To determine fish diameters, we first converted lengths to girth (i.e., circumference) using published length/girth ratios (Jones et al. 1999). We then converted girths to diameters, assuming a perfect circle, using the formula d = c/pi. Fish with d less than or equal to 6.35 mm (i.e., less than 34.8 mm total length) were then removed from the dataset prior to analysis.

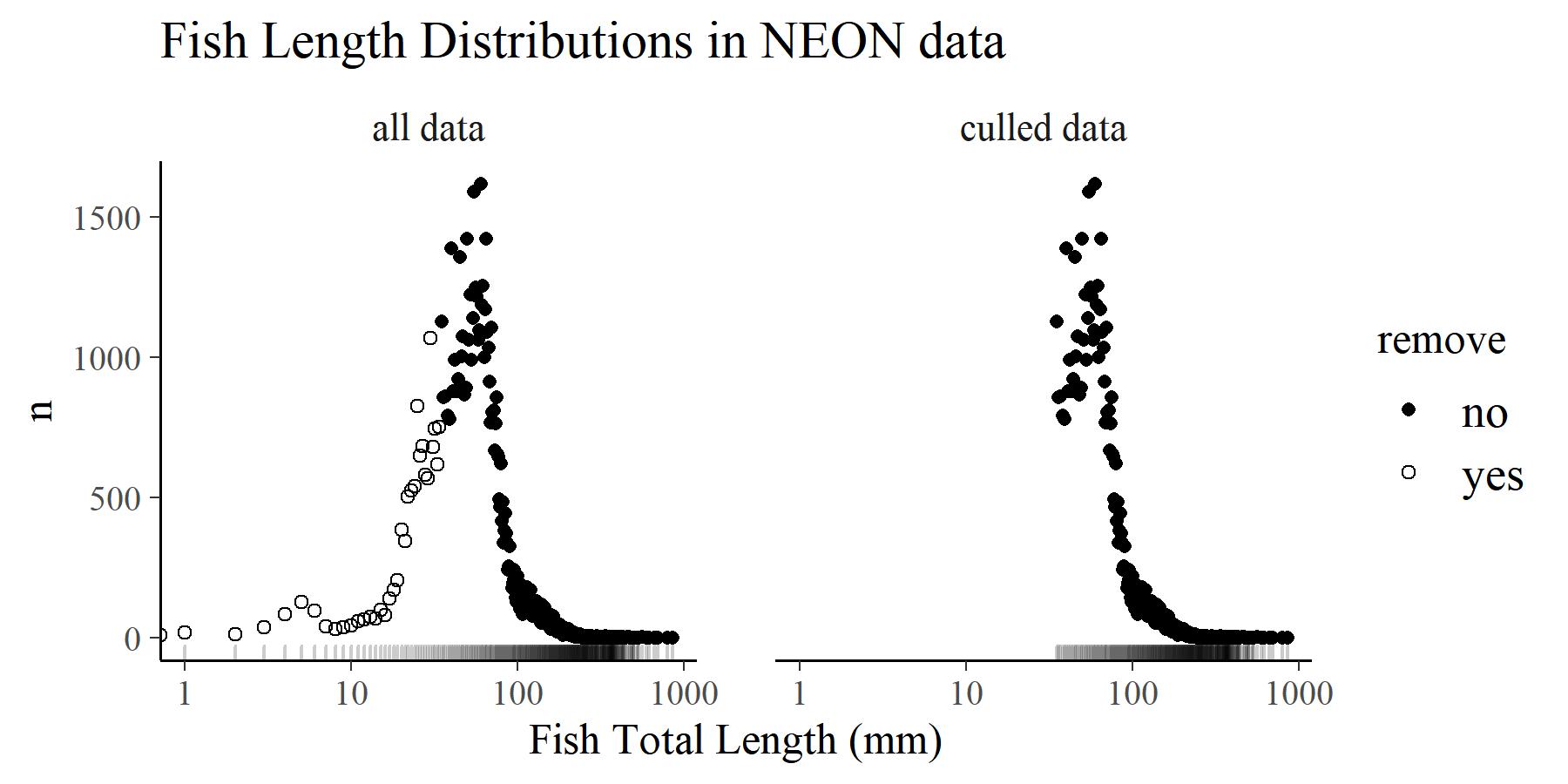


Figure X. Distributions of ~70,000 individual fish lengths. After converting lengths to diameters, we excluded fish that were less than 2 times the width of the mesh diameters to limit bias against small fish. Rugs on the x-axis show the distribution of data in each dataset.

After measuring the first 50 or so fish, NEON then bulk counts the rest and reports those counts per pass in the table *fsh\_bulkCount*. For example, if 100 Rainbow Trout were collected on the first pass, then *fsh\_bulkCount* would indicate 50 fish with the other 50 indicated by the total count in *fsh\_perFish*. Adding these totals gives the total number of fish per reach per pass per taxon. We used these data to estimate the total population size of each species using the CarleStrub formula in the *removal()* function of the FSA package (CITE). This resulted in estimates of fish abundance per m2. We converted these estimates to fish *size* per m2 by sampling sizes from the fish size measures. To do this, we first converted measures from per\_m2 to 1000\*per\_m2 so that abundance estimates were in integers. We then sampled individual body sizes (mg dry mass) with replacement from the *fsh\_perFish* table relative to the number of individuals estimated from the 3-pass removal. For example, if the *removal()* function estimated 296 individuals of Speckled Dace in a reach, then we sampled 296 body sizes from the Speckled dace body size measurements (at that site/date/reach combination). Since only 50 individuals were measured, that meant that some body sizes are repeated. From this list of body sizes, we obtained a list of counts per body size, which we then converted to body size per m2. These estimates were then combined with macroinvertebrate estimates to generate a list of body sizes (mg dry mass) and their abundance (no/m2) for both fish and macroinvertebrates.